

UNIVERSIDAD DE CIENCIAS MÉDICAS
HOLGUÍN



MAESTRIA EN EDUCACION MÉDICA SUPERIOR

Titulo: Calidad del examen final de la asignatura Metodología de la Investigación y Estadística. Carrera de Medicina. Curso 2010-2011.

Autora: Lic. Gemma M. Ortiz Romero.

Tutor: Dr. Pedro A. Díaz Rojas.

Trabajo para optar por el Título de Master en Educación Médica Superior

2012

“Año 54 de la Revolución”

Dedicatoria:

A mis padres, por su dedicación y sabios consejos durante toda la vida.

A mis hijos, como legado al esfuerzo y la tenacidad para el crecimiento personal y profesional.

A mi esposo, por su comprensión y ayuda incondicional en todo momento.

Agradecimientos.

A mi tutor, el Dr. Pedro Díaz Rojas por su acertada y eficiente guía en el desarrollo de esta investigación y su estímulo permanente.

A la doctora Magalys Moreno Montañez por sus oportunos consejos.

Al colectivo de profesores de la maestría, por su contribución a mi formación.

A mis compañeros del Departamento de Informática Médica por su apoyo, en especial a la Jefa de departamento, María Emilia Rodríguez por su constante apoyo espiritual y material en el desarrollo de este trabajo.

A todas aquellas personas que hicieron posible la realización de la presente investigación.

Resumen.

La evaluación del progreso de los estudiantes es un asunto polémico en cualquier modelo pedagógico y un medio para mejorar la calidad del proceso docente educativo. Se realizó una investigación en el campo de la evaluación del aprendizaje, con el objetivo de caracterizar el examen final de la asignatura Metodología y Estadística aplicado a los estudiantes de segundo año de la carrera de Medicina en la Facultad de Ciencias Médicas de Holguín en el curso 2010 - 2011. Del universo formado por todos los exámenes finales ordinarios de la asignatura en la provincia Holguín, se seleccionó una muestra intencionada de las 276 pruebas de los estudiantes que se presentaron a esta convocatoria en el municipio Holguín. Los métodos teóricos, empíricos y procedimientos estadísticos empleados permitieron la triangulación de la información. Se entrevistó a 6 expertos, se analizó la correspondencia del fondo de tiempo y la cantidad de ítems explorados en el examen, el índice de dificultad, el coeficiente de correlación del punto biserial y el Alfa de Cronbach. Los resultados evidenciaron predominio de las preguntas de tipo ensayo de respuestas cortas, la no correspondencia entre el tiempo dedicado a cada tema y la cantidad de ítems explorados en el examen, escasa representación del tema Metodología de la Investigación, el examen clasificó con una dificultad media, excelente discriminación de todas las preguntas aunque no de todos los ítems y una confiabilidad aceptable.

Índice.

Introducción	1
Marco teórico	7
Objetivos	43
Diseño metodológico	44
Análisis y discusión de los resultados	50
Conclusiones	65
Recomendaciones	66
Referencias bibliográficas	67
Anexos	i

Introducción.

La Educación Superior a lo largo de su historia ha dado sobradas pruebas del grado de identificación e interacción con los principales problemas de su territorio y del trabajo que realiza en la solución de éstos. De ello depende la efectividad de los procesos educativos, así como de su capacidad para transformarse y propiciar el cambio en el progreso de la sociedad.

Según el alcance y el ritmo de estas transformaciones, la sociedad tiende cada vez más a fundarse en el conocimiento; razón para que la Educación Superior y la investigación formen hoy en día parte fundamental del desarrollo cultural, socioeconómico y ecológicamente sostenible de los individuos, las comunidades y las naciones. En Latinoamérica la década del sesenta marcó el paso de una educación superior de carácter elitista, a otra de carácter masivo en la que Cuba es un buen ejemplo.¹

En la Educación Médica Superior cubana también se producen transformaciones, observadas desde el inicio del proceso revolucionario cuando a partir de la década del 60, se puso en marcha el Policlínico Integral, que se modificó en los años 70 para dar paso al Policlínico Comunitario. En la década de los 80, se implanta al plan del médico y enfermera de la familia hasta las transformaciones más recientes, el Policlínico Universitario y el Programa de Formación de Médicos Latinoamericanos.² Basadas en el principio de formación de los estudiantes en los escenarios reales de actuación, para permitir así la integración de los conocimientos teóricos con los problemas de salud de la comunidad.³

Estas condiciones hacen más complejo el proceso formativo, y como plantea Fidel Ilizástigui “Educar al hombre como médico para que actúe toda su vida como buen ciudadano y excelente profesional, comprometido con la salud poblacional, es tarea extremadamente difícil y compleja”.⁴

El desarrollo de la educación como una necesidad que imponen los diferentes cambios sociales, entre otros factores, condiciona su perfeccionamiento y al mismo tiempo el de la evaluación, que es uno de los asuntos más polémicos en cualquier modelo pedagógico⁵ y uno de los medios para mejorar la calidad del proceso docente educativo.

La evaluación del progreso de los estudiantes a lo largo de la carrera es un dilema presente en la formación inicial del médico. Durante años estuvo dirigida solo a medir los conocimientos que los estudiantes obtenían en cada disciplina académica. Desde mediados de la década del 70 comienza a adquirir un peso importante la evaluación de las habilidades clínicas de los estudiantes y su competencia para la solución de los problemas de salud de los pacientes.⁶

En diferentes espacios para debatir sobre Educación Médica tales como: la Cumbre Mundial de Educación Médica de Edimburgo (1993), en el Encuentro Continental de Educación Médica de Punta del Este (1994), y en el Encuentro Mundial de Educación Médica de Santa Fe de Bogotá (1995) se produce una seria reflexión sobre cómo la gestión universitaria puede impactar en la calidad de la formación y desarrollo de los recursos humanos en salud. Entre las medidas propuestas se encuentra el perfeccionamiento de los sistemas evaluativos del pregrado y posgrado.⁶

Desde el curso 2004-2005 se inicia en 75 centros del país el Proyecto Policlínico Universitario con la participación de 1 000 estudiantes de primer año de medicina. Estas unidades de la Atención Primaria de Salud fueron dotadas de la tecnología necesaria para el nuevo modelo pedagógico, y se convierten en el escenario principal del proceso de aprendizaje⁷ donde se amplían los espacios para la evaluación permanente de los estudiantes.

El gran valor de la evaluación radica entre otros aspectos, en ser "instrumento de investigación didáctica: comprobar hipótesis de acción metodológica para ir acumulando recursos de una eficacia comprobada en la acción, e ir engrosando de esta manera el apartado de técnica pedagógica fundamentada científicamente". La evaluación, por tanto, presenta este doble carácter, el ser un instrumento de investigación didáctica y a la vez, un medio pedagógico que proporciona una información para ayudar al estudiante a progresar en su aprendizaje.⁸

En este sentido, se destaca que la eficacia de la evaluación habrá de admitirse desde estas dos perspectivas, la evaluación del proceso educativo y la evaluación del aprendizaje del estudiante, pero el objetivo de la evaluación va más allá de estos aspectos.⁸ No se trata de tomar decisiones ni sobre el proceso ni sobre el

producto, sino que un colectivo tome conciencia de su situación para plantear mejoras en la enseñanza.

A partir del análisis anterior, se revela la necesidad de profundizar en las relaciones que se establecen, en el contexto del proceso docente de la asignatura Metodología de la Investigación y Estadística en la carrera de Medicina, entre el desarrollo del proceso en la referida asignatura y los resultados alcanzados en los instrumentos evaluativos que se aplican.

La asignatura Metodología de la Investigación y Estadística posibilita desarrollar la capacidad organizativa y el hábito de proceder reflexivamente en el enfrentamiento de los problemas relacionados con el tratamiento metodológico y estadístico de la información médica, a la vez que vincula los contenidos de la asignatura con la práctica de la profesión y la vida social en los diferentes niveles de atención de salud.⁹

Entre los objetivos terminales del Médico General plasmado en el perfil del egresado, se encuentra que este profesional debe ser capaz de aplicar el método científico al diagnóstico y solución de los problemas de salud del individuo, la familia y la comunidad, a la búsqueda, evaluación y aplicación de la información científica técnica relacionada con la salud humana, a la búsqueda y recolección activa de la información y su análisis estadístico, tanto en el ejercicio cotidiano de su profesión, como en su participación en la ejecución de investigaciones de carácter regional o nacional en su área de trabajo,¹⁰ en este aspecto la asignatura Metodología de la Investigación y Estadística es de vital importancia para los profesionales del sector de la salud.

La autora a través de su experiencia como docente y al evaluar a los estudiantes, participar en tribunales de calificación de exámenes de la asignatura, la revisión bibliográfica, el criterio de otros miembros del colectivo de trabajadores, pudo constatar que durante el proceso evaluativo de la asignatura en la Facultad de Ciencias Médicas de Holguín se presentan insuficiencias, que limitan el cumplimiento de los objetivos planteados con anterioridad, entre las que se destacan: los resultados de los diferentes instrumentos evaluativos que se aplican no se corresponden con los resultados esperados, con una marcada diferencia en el examen final; gran cantidad de estudiantes suspensos o con baja calidad en las

evaluaciones finales y parciales, que no están en correspondencia aparente con la labor desarrollada por el colectivo pedagógico; insatisfacción de los estudiantes por los resultados alcanzados; no se cuenta con una valoración científica de la calidad del examen final en esta asignatura.

A partir de estas insuficiencias se formulan las siguientes interrogantes ¿por qué los resultados de las evaluaciones finales de los estudiantes son deficientes?; ¿qué nivel de confiabilidad poseen los instrumentos aplicados?; ¿qué nivel de dificultad y discriminación tienen las preguntas y los ítems de los instrumentos aplicados?

La construcción de forma empírica de los instrumentos para evaluar el aprendizaje de nuestros estudiantes ocasiona problemas con su calidad, con relativa frecuencia sus resultados no miden el cumplimiento de los objetivos educacionales propuestos y las calificaciones otorgadas no expresan de forma real y objetiva su nivel de aprovechamiento como para generalizarlos y determinar el nivel alcanzado en su competencia profesional, en correspondencia con la escala de calificación empleada.⁶

El uso de los instrumentos por parte de quienes deben llevar a cabo la mejora de la docencia, puede favorecer su repercusión en el incremento del nivel de calidad de la misma, si se dan las condiciones adecuadas.¹¹ Es necesario que el sistema de evaluación sea lo más sólido posible, y aumente la importancia de que los instrumentos a emplear en los controles tengan la calidad requerida y se pongan a prueba.¹²

Para poder evaluar calidad es necesario que se realice un análisis sistemático e interpretación de los resultados obtenidos, mediante la aplicación de diferentes instrumentos de medición, con la finalidad de adoptar las mejores decisiones posibles en cada caso o situación.²

El análisis de los instrumentos evaluativos y de sus resultados toman auge a nivel mundial en la década de los años 80 del pasado siglo en centros de Educación Médica en Escocia y Toronto, de donde surge la idea de celebrar conferencias en Ottawa como un marco adecuado para la discusión de este tema.¹³

En 1985, la Asociación Americana de Psicología, en colaboración con la Asociación Americana de Investigación Educativa y el Consejo Nacional para la

Medición en Educación, publicó el manual “Estándares para las pruebas educativas y psicológicas”, documento que enfatiza la importancia de tomar en cuenta el bienestar de las personas que hacen una prueba y evitar el mal uso de los instrumentos de evaluación. Mientras que, en países desarrollados, es obligatorio que estos criterios de calidad se satisfagan, en otros países se asumieron con posterioridad.¹⁴

En México existen estudios que tratan sobre el nivel de dificultad y de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA), para el ingreso de bachilleres a la educación superior.¹⁵ Otros estudios describen el nivel de facilidad y poder de discriminación de Inferencia Estadística en Medicina Veterinaria y Zootecnia.¹⁶

En España se realizaron estudios sobre Análisis de los ítems de las evaluaciones autonómicas de diagnóstico,¹⁷ sobre Matemática. En Bolivia se estudió el Nivel de dificultad y poder de discriminación del tercer y quinto examen parcial de la cátedra de cito-histología carrera de medicina de la UMSA.¹⁴

Otros estudios fueron realizados en la República Bolivariana de Venezuela por integrantes de la colaboración médica cubana en sus trabajos de culminación de maestría en Educación Médica, centrados en el análisis de instrumentos aplicados en el examen estatal escrito de la especialidad de Medicina General Integral² y en diferentes evaluaciones de la asignatura Morfofisiología Humana en la carrera de Medicina.^{18- 20.}

En Cuba se encontraron diferentes trabajos realizados en Ciencias Médicas, entre los que se destacan los realizados en Licenciatura en Enfermería²¹ y en Facultades de Medicina de otras provincias.^{12, 22.} Los resultados obtenidos aportan información sobre la calidad de los exámenes diagnósticos o finales de la asignatura Morfofisiología Humana en diferentes semestres de la carrera, en exámenes estatales o de la especialidad de MGI, con el uso de diferentes indicadores.

La autora de este trabajo considera importante destacar el estudio realizado por Enrique Santiesteban (2007) para optar por el título de Máster en Educación Médica, en el que realiza una propuesta de metodología para el análisis de los resultados de los exámenes finales de las asignaturas de la disciplina

Morfofisiología Humana del Programa Nacional de Formación de Medicina Integral Comunitaria en la República Bolivariana de Venezuela,²³ en la que considera proponer además del índice de dificultad; el análisis de la correspondencia entre el fondo de tiempo dedicado a cada una de las temáticas de la asignatura y el número de incisos en que se explora cada una de ellas en el examen.

En la revisión documental realizada no se encontró estudios en la Universidad de Ciencias Médicas de Holguín durante el pregrado ni el posgrado, evidencia que en este centro solo se aplican instrumentos de evaluación contruidos de forma empírica, situación común para la asignatura Metodología de la Investigación y Estadística. En estos momentos, la Facultad de Ciencias Médicas "Mariana Grajales Coello" se encuentra inmersa en el proceso de acreditación de la carrera y constituye un componente esencial la evaluación de todos los procesos que se ejecutan.

Por todo lo expuesto, el **problema científico** que sustenta la presente investigación expresa que no existe una caracterización sobre bases científicas de la calidad de los instrumentos evaluativos en la asignatura Metodología de la Investigación y Estadística.

Las valoraciones teóricas y metodológicas sobre la evaluación del aprendizaje sirven de referente a la asignatura Metodología de la Investigación y Estadística. Por primera vez se realiza la caracterización científica de la calidad del examen final a la referida asignatura que permite obtener elementos necesarios para el perfeccionamiento de los instrumentos de evaluación que se aplican.

A partir de este estudio se podrán implementar estrategias metodológicas en el departamento de Informática Médica para elevar la calidad de los instrumentos de evaluación, lo que favorece la calidad del proceso docente educativo y repercute en el nivel de competencias de los profesionales que egresen.

Marco teórico.

La Educación Superior Cubana pone todo su empeño en fortalecer el proceso de formación integral de los estudiantes universitarios, abordado por Ilizastigui²⁴ como "... la pretensión de centrar el quehacer de las universidades en la formación de valores en los profesionales de forma más plena, dotándolo de cualidades de alto significado humano, capaces de comprender la necesidad de poner sus conocimientos al servicio de la sociedad en lugar de utilizarlos sólo para su beneficio personal. Implica también la necesidad de lograr un profesional creativo, independiente, preparado para asumir su auto educación durante toda la vida..."

La universidad médica.

La formación de profesionales con una alta preparación científica, cultural y humanista, constituye una prioridad para todos los Ministerios de Salud del mundo y de Cuba, desde esta perspectiva la Educación Médica Superior tiene el reto de formar un profesional de perfil amplio con profundos conocimientos que den respuesta a los problemas de salud que se presenten una vez graduado.²⁴

Unido a este reto, el desafío actual que en el campo de la educación médica tiene nuestro país, determinado por la necesidad de fundamentar científicamente todas las transformaciones realizadas y las nuevas modalidades surgidas a la luz del modelo cubano de formación médica, frente al fenómeno progresivo de la universalización de la enseñanza universitaria, provoca que se hayan multiplicado las necesidades de realización de investigaciones educacionales con el objetivo de evaluar y perfeccionar los procesos formativos.²⁵

En el desarrollo de estos procesos la universidad médica ha puesto en evidencia la necesidad de lograr un espacio mayor dentro del proyecto curricular al aprendizaje de los estudiantes en los escenarios propios de la atención primaria en salud y obedece a la necesidad de desarrollar habilidades, hábitos y convicciones acerca del quehacer propio de este Sistema.⁷

Desde el surgimiento de la carrera de Medicina en nuestro país en el año 1726 hasta la actualidad se aplican varios planes de estudios, que se perfeccionan según el momento histórico social. El plan actual se pone en práctica a

partir de 1985 y constituye una positiva respuesta para el ampliado proceso de formación de médicos, en las condiciones referidas a inicios de los 80, con aportes innovadores en su concepción y ejecución. Un significado especial tiene su proyección a la Atención Primaria de Salud, acorde con el modelo de atención del médico y la enfermera de familia, de acuerdo con la estrategia del Comandante en Jefe Fidel Castro Ruz en el campo de la salud, año en el cual surge como especialidad la Medicina General Integral.²⁴

Este plan de estudio sufre transformaciones para su continuo perfeccionamiento según los requerimientos del sistema de salud y en busca de un mejor desarrollo del proceso docente educativo, en respuesta a los avances científicos y las tendencias de la Educación Médica.

La última de estas transformaciones se enmarca en el año 2010, donde se precisa que el tipo de médico a egresar debe ser un Médico General, se define la atención médica integral, la docente-educativa, la investigación, la administración y las especiales como las cinco funciones básicas de este profesional, definida como función rectora la Atención Médica Integral, se precisa que la función de investigación debe tener una proyección general, no quedar circunscrita a las biomédicas y también se muestran cambios en la disciplina Informática Médica.

La disciplina Informática Médica.

La Estadística ha devenido en una asignatura de obligatorio estudio en la educación universitaria. Tal es así que todas las carreras universitarias tienen en su currículo un programa de esta asignatura, la necesidad de analizar grandes cantidades de datos recopilados de procesos experimentales para la obtención de conclusiones y la toma de decisiones sobre los fenómenos objeto de estudio hacen de las herramientas estadísticas un recurso de importancia trascendental.²⁶

El médico debe estar preparado para aplicar las tecnologías de la información, los métodos para la recolección, procesamiento, análisis y presentación de datos en su desempeño laboral y en su educación continuada en el postgrado. Debe ser capaz de evaluar continuamente los factores que inciden de manera negativa en el correcto desenvolvimiento de su labor profesional, y está llamado a diseñar estrategias que permitan resolver los problemas de salud que se dan en el entorno que desarrolla su profesión, a partir de la investigación científica. La

Informática Médica como disciplina del plan de estudio de la carrera de Medicina juega un rol esencial en la formación de este profesional.

La Informática Médica emerge de la intersección de las ciencias de la salud y las del procesamiento de la información.²⁷ Se ocupa de los recursos, los dispositivos y los métodos necesarios para optimizar la adquisición, almacenamiento, recuperación y utilización de la información en salud y biomedicina.

Entre los conceptos que se asocian a la Informática Médica está el que plantea Shortliffe, E. H. (1984), "La ciencia de la información médica es la que utiliza sistemas como herramientas analíticas para el desarrollo de procedimientos que permitan la manipulación, el control de procesos, la toma de decisiones y el análisis científico del conocimiento médico".²⁸

Como disciplina docente, comprende los recursos de las tecnologías de la información y la comunicación, la implementación de estrategias científicas y metodológicas eficientes para la solución de problemas integradores profesionales de las diferentes esferas de las Ciencias Médicas.²⁹

La Reforma Universitaria que se consolida en 1962, señala el momento del inicio de cambios en la enseñanza de la salud pública y los contenidos médico-sociales en la carrera de Medicina en Cuba. En el plan de estudios que se aprueba emerge la asignatura Estadística de Salud, trataba contenidos de estadística y metodología de la investigación.³⁰

En 1970 se incluye en el plan de estudios de la carrera de Medicina para el ciclo básico, la asignatura El Hombre y su Medio, con un módulo de conferencias que impartía el tema Estadística Básica y Metodología de la Investigación. En el año 1974 debido a la necesidad de especializar profesionales de la salud en Bioestadística, se aprueba en el Ministerio de Salud Pública esta especialidad con un programa de estudio totalmente estructurado.³¹

Entre los años 1982 y 1983, la microcomputadora como avance científico y como herramienta de trabajo ya era un hecho consolidado a nivel mundial que ganaba terreno en la sociedad cubana, el Ministerio de Educación Superior (MES) orientó la creación de los Planes Directores de Computación en todas las especialidades de la enseñanza universitaria.^{27, 32.}

En Ciencias Médicas ocurrió algo muy peculiar. En el plan de estudio de Medicina,

carrera que se tomó siempre como referencia para las adecuaciones, no había espacio para la asignatura Computación y no se disponía de personal docente capacitado a tales fines.²⁷

Tras un análisis minucioso de la situación, se considera que los únicos profesores que pueden asumir la docencia de estos contenidos eran los de la asignatura Bioestadística por la preparación matemática que poseían.

Cuando el MES decide incorporar la Computación como asignatura en todas las especialidades universitarias, es Ciencias Médicas la última en ejecutar oficialmente dicha medida en el curso 1985-1986. En la asignatura Bioestadística se incluyen los contenidos de Computación; cambia el nombre a Bioestadística y Computación y se imparte en el cuarto semestre de la carrera.^{27, 32.}

En este período la evaluación es teórica y se centra fundamentalmente en los temas de Bioestadística.

Con la creación de los laboratorios de computación en 1990 cambia el programa, en la práctica social el profesional de la salud no está llamado a programar sobre el equipo de cómputo, sino a explotarlo como herramienta de trabajo.

En 1991 con la entrada de nuevas computadoras nuevamente sufre transformaciones el programa de la asignatura. Se introduce el enfoque usuario de la computación con el estudio de aplicaciones típicas de los sistemas personales como procesadores de textos y gestores de bases de datos, además de graficadores y paquetes estadísticos que constituyen la máxima expresión de integración de la Bioestadística y la Computación.^{27, 32.} En estos momentos la mayoría de los estudiantes utilizaban la computadora solo en esta asignatura.

El advenimiento de la sociedad de la información es consecuencia del desarrollo científico - tecnológico, especialmente en las áreas de la Informática y las telecomunicaciones, hizo posible la aparición de tecnologías avanzadas denominadas Nuevas Tecnologías de la Información y trajo como consecuencia que partir de 1997, se sustituyeran los equipos de los laboratorios docentes por máquinas Pentium hecho que conllevó a una revitalización de la asignatura Bioestadística y Computación, surge así la disciplina Informática Médica que incluye la asignatura Informática Médica I e Informática Médica II.^{27, 33.}

La Informática Médica I se ubica en el cuarto semestre de la carrera con una

duración de 72 horas lectivas, tiene como sistema de conocimientos los temas relacionados con introducción a la manipulación de la información; aspectos relacionados con el tratamiento de la información; metodología de la investigación y estadística descriptiva. La Informática Médica II ubicada en el sexto semestre de la carrera con una duración de 20 horas lectivas, tiene como sistema de conocimientos un solo tema relacionado con la estadística inferencial.^{33, 34.}

Se puede apreciar que en la asignatura Informática Médica I, se unen los temas de Informática con los de Bioestadística con una notable diferencia en el total de horas clases de cada uno.

Con el surgimiento y desarrollo de Infomed como la red telemática de salud pública, permite el enlace de todas las Universidades de Ciencias Médicas del país, con acceso a importantes bases de datos internacionales, así como se brindan varios servicios de Internet. El funcionamiento estable de esta red fortalece el uso de la información científico-investigativa y postgraduada.

En el curso 2000 – 2001, la asignatura Informática Médica I se imparte en el segundo semestre de la carrera con una duración de 60 horas lectivas, incluye los contenidos de Computación. La asignatura Informática Médica II se imparte en el cuarto semestre de la carrera con 60 horas lectivas, incluye los temas metodología de la investigación; estadística descriptiva; elementos de demografía y estadísticas sanitarias e inferencia estadística.⁹

El sistema de evaluación se modifica al contar con las computadoras, se aplican exámenes prácticos con el uso de sistemas estadísticos para el procesamiento de los datos, pero no se realiza la interpretación médica de la información obtenida.

En este período se actualizan los contenidos de computación para la asignatura Informática Médica I (Sistema Operativo Windows) y se incorporaron los temas de Redes y Correo Electrónico.

A partir del año 2003 con la implementación de los Programas de la Revolución, que incluye como priorización la informatización de los servicios, se inicia un amplio proceso, orientado en primer lugar a la superación y desarrollo profesional que, a su vez, se extenderá con la automatización de los servicios médicos, la investigación, la información científico-técnica y el apoyo en la toma de decisiones.³⁴

El sistema de evaluación se mantiene igual y los instrumentos utilizados se limitan a ejercicios descontextualizados de mínima complejidad, limitan la creatividad, la integralidad y la relación interdisciplinaria y por tanto su vínculo con la práctica médica.³¹

En el curso 2004-2005 se crea el Proyecto Policlínico Universitario, que incorporó un nuevo modelo pedagógico con la aplicación el principio de la universalización de la enseñanza y el traslado del escenario principal de la docencia a las áreas de atención primaria de salud (policlínicos y consultorios), la utilización de la tecnología educativa y la amplia utilización de las computadoras y el software educativo en el proceso pedagógico. El proceso de enseñanza aprendizaje se centraba en las conferencias impartidas a través del video y luego se realizaban las guías de estudios en la computadora.³⁴

La utilización de estas tecnologías en la Educación Médica han introducido en este campo nuevos paradigmas educativos, como son la educación centrada en el estudiante, el autoaprendizaje y la gestión del conocimiento, modifican también el papel histórico de los profesores, pues se convierten en facilitadores, moduladores y moderadores del proceso.³⁵

La asignatura Informática I pasa al primer semestre de la carrera, por la vital importancia que tiene la computadora como elemento básico en el quehacer cotidiano de estudiantes y profesores en este nuevo proyecto de formación de profesionales de la salud. La asignatura Informática Médica II se mantiene con los mismos temas de estudio.

En consecuencia, las dos asignaturas de la disciplina Informática Médica experimentaron algunos cambios en cuanto a las formas de organización del proceso, la planificación en el tiempo de las actividades docentes y los métodos y medios de enseñanza-aprendizaje.³⁶

En el policlínico universitario se realiza mayor cantidad de evaluaciones teórico prácticas en la asignatura Informática Médica II por la gran disponibilidad de recursos informáticos, pero es insuficiente interrelación entre los temas de estudio.

A partir del curso 2007-2008 quedan definidos tres modelos de formación para Medicina: el modelo del Policlínico Universitario, cuyos escenarios docentes son

los policlínicos y hospitales; el modelo Tradicional que se desarrolla en las propias facultades y el Nuevo modelo de formación del médico latinoamericano que se implementa en algunos escenarios.

Se plantea la necesaria aplicación práctica de los contenidos del tema de Metodología de la Investigación, el tiempo dedicado al mismo resulta insuficiente para la formación de habilidades investigativas generales como la formulación de problemas, objetivos e hipótesis a partir del vínculo con la práctica.³⁶

En el año 2010 se realiza el perfeccionamiento del plan de estudio de la carrera de Medicina y se producen cambios en la disciplina Informática Médica, en el caso de la Informática Médica I se propone el tránsito al software libre y la asignatura Informática Médica II cambia su nombre, ahora denominada Metodología de la Investigación y Estadística y ocurren cambios en la ubicación de los temas, aumenta las horas clases al tema de Metodología de la Investigación y se incluyen elementos de Metodología cualitativa.¹⁰

En las evaluaciones se logra una mayor vinculación entre los temas de la asignatura, con los problemas reales que enfrenta el médico en la atención primaria de salud y con la preparación del estudiante en su función investigativa, la que aún es insuficiente.

La clase de “Metodología de la Investigación y Estadística” debe centrarse en la participación activa de los estudiantes, en el proceso, desde la formulación de preguntas clave, la recogida, organización y representación de datos, análisis de los mismos y elaboración de conjeturas, hasta la comunicación de la información obtenida de una manera clara y precisa.³¹ El sistema de evaluación se mantiene con evaluaciones sistemáticas, un seminario, dos trabajos de controles, una prueba intrasemestral y una prueba final.

La inclusión de esta asignatura en el diseño curricular de la carrera de Medicina se fundamenta a partir de la necesidad de conocimientos de los profesionales de la salud sobre métodos cuantitativos a partir de las posibilidades que brindan para la solución de múltiples problemas en el área de las ciencias médicas³⁶, y elementos de Metodología de la Investigación que le permitan, con el apoyo de los recursos informáticos, desarrollar investigaciones científicas para resolver problemas de salud.

El proceso docente educativo.

“La didáctica es la ciencia que estudia como objeto el proceso docente educativo (PDE) dirigido a resolver problemática que se plantea a la escuela: La preparación del hombre para la vida, pero de un modo sistemático y eficiente”. El objeto de estudio de la didáctica es el proceso docente educativo, entendiéndose por éste “el proceso eficaz y eficiente que de respuesta al encargo social”.³⁷

La eficacia del proceso docente educativo mide los resultados alcanzados en función de los objetivos propuestos. Su eficiencia se expresa en graduados capaces de cumplir con la máxima calidad, el encargo que le plantea la sociedad. Mayor eficacia se logra en la medida que las distintas etapas necesarias para arribar a esos objetivos, se cumplen de manera organizada y ordenada sobre la base de su prioridad e importancia.

Para autores como María Rosa Milán, Homero Fuentes y colaboradores, el PDE se concibe como el proceso que de modo consiente se desarrolla en las instituciones, a través de las relaciones de carácter social que se establecen entre sus participantes, con el propósito de educar, instruir y desarrollar a los futuros profesionales, sistematiza y desarrolla de forma planificada y organizada la cultura acumulada por la humanidad que dan respuestas a las demandas de la sociedad.³⁸

La autora Olga de Landaluce en su libro Pedagogía, dedicado a estudiantes y profesores de Tecnología de la Salud plantea "La práctica histórico social demuestra que la formación de las nuevas generaciones de acuerdo con las aspiraciones de la sociedad se produce, fundamentalmente en el proceso docente educativo".³⁹ Indica que la institución docente y el proceso que en ella se desarrolla existen para satisfacer una necesidad social y la formación de profesionales competentes: encargo social que plantea la sociedad a la universidad médica cubana.

El Reglamento de Trabajo Docente Metodológico del Ministerio de Educación Superior de Cuba refiere: la labor que, apoyados en la Didáctica, realizan los sujetos que intervienen en el PDE, con el propósito de alcanzar óptimos resultados en dicho proceso, jerarquizar la labor educativa desde la instrucción, para satisfacer los objetivos formulados en el plan de estudio, es la esencia del

trabajo metodológico. Se concreta en el desarrollo con calidad del PDE, para el logro de una adecuada integración de las clases con la actividad laboral e investigativa y en las tareas de alto impacto social.⁴⁰

Para Álvarez de Zayas³⁷ el trabajo metodológico es el proceso de dirección del PDE, en el cual se desarrolla tanto la planificación y organización del proceso como su regulación y control.

En los centros de educación superior, el trabajo metodológico se realiza de forma individual y colectiva; la individual es la labor de autopreparación que realiza el profesor y constituye la premisa fundamental para que resulte efectivo el trabajo metodológico que realiza el profesor de forma colectiva. La colectiva, tiene como rasgo esencial el enfoque en sistema y se lleva a cabo en cada uno de los niveles organizativos del proceso docente educativo, como vía para su perfeccionamiento en cada nivel.⁴⁰

Entre los subsistemas o niveles organizativos principales para el trabajo metodológico se encuentra el colectivo de carrera, colectivo de año, colectivo de disciplina e interdisciplinarios (en casos necesarios) y el colectivo de asignatura.⁴⁰

El colectivo de disciplina responde por el trabajo metodológico en este nivel organizativo. Agrupa a los jefes de colectivo de las asignaturas de la sede central y los coordinadores de la disciplina en las sedes universitarias. Su propósito fundamental es lograr el cumplimiento con calidad de los objetivos generales de la disciplina.

Por otra parte el colectivo de asignatura responde por el trabajo metodológico en este nivel organizativo. Agrupa a los profesores que desarrollan la asignatura. Su propósito fundamental es lograr el cumplimiento con calidad de los objetivos generales de la asignatura, en estrecho vínculo con los objetivos de la disciplina y del año en el cual se imparte.

Las formas de trabajo metodológico en la Educación Superior son el trabajo docente metodológico y el científico metodológico que de forma escalonada se ejecutan en cada subsistema según los principales problemas detectados en la conducción del proceso docente.⁴⁰

Como consecuencia del perfeccionamiento del trabajo metodológico de los claustros durante varios años la evaluación ha transitado desde un enfoque

acumulativo-cuantitativo, propio de etapas anteriores de desarrollo a uno cualitativo-integrador, centrado en el desempeño de los estudiantes.⁵

La autora del presente trabajo considera importante profundizar en el desarrollo del trabajo metodológico dirigido a perfeccionar el sistema de evaluación del aprendizaje de la asignatura, como elemento esencial para elevar la maestría pedagógica de los profesores y consolidar la eficiencia en el trabajo de los diferentes colectivos de asignaturas y disciplinas, núcleo básico donde radica el éxito del proceso de enseñanza-aprendizaje.

Concuera a su vez, con el criterio de otros autores al plantear que en el PDE juegan un rol decisivo los componentes personales de dicho proceso, constituido por profesores y estudiantes.^{5, 6, 37, 40.}

En la concepción del PDE tradicional, se considera el estudiante como objeto, donde recae la actividad del profesor para que aprenda, el aprendizaje es el resultado y el proceso que dirige el profesor es la enseñanza, por tanto el estudiante es un ente pasivo en dicho proceso.³⁹

Aunque los diferentes modelos que sirven de base a la práctica escolar en la actualidad han tenido en el centro la comprensión de un sujeto activo, reflexivo, con el propósito de minimizar las actividades pasivas de los alumnos, aún persisten algunos reflejos de una enseñanza tradicional.³⁹

Este proceso debe concebirse como una actividad íntegra, que según las concepciones actuales, se centra en el papel protagónico del estudiante. Es el marco donde se lleva a cabo la construcción de conocimientos, habilidades, valores, sentimientos, actitudes y normas de comportamiento por parte de los estudiantes bajo la oportuna y efectiva planificación, dirección, ejecución y control del aprendizaje de los educandos por parte del profesor,^{37, 41} debe caracterizarse por el énfasis en el aprendizaje independiente del estudiante, quien tiene el rol principal, bajo la guía del profesor y apoyarse en la utilización de métodos activos de aprendizaje y de recursos informáticos.

En la actualidad en Cuba, el proceso de aprendizaje que se asume tiene como sustento la teoría pedagógica del enfoque histórico - cultural. Un elemento esencial lo constituye la consideración de la enseñanza como guía del desarrollo, los niveles que alcanza el estudiante como parte de su aprendizaje estarán

mediados por la actividad y la comunicación,⁴² ambas constituyen agentes mediadores entre el estudiante y la experiencia cultural que va a asimilar.

En este contexto cobra vital importancia el aprendizaje productivo, donde el estudiante adopte una posición activa, aporte sus criterios, vía de solución de los problemas y se involucre en el control de sus propias acciones de aprendizaje.

La autora comparte los criterios sobre aprendizaje dados por Héctor Valdés y Francisco Laos al expresar: el proceso de aprendizaje que se desarrolla en el grupo de estudiantes encuentra en el profesor su mediador esencial,⁴² en esta concepción se le concede un gran valor a los procesos de dirección y orientación que estructura el docente sobre la base de una intención educativa, expresada en el fin de los objetivos a alcanzar en todos los estudiantes a partir de sus potencialidades particulares, lo que expresa el par dialéctico de calidad y masividad.

De igual forma coincide con lo planteado por Olga de Landaluce "el profesor en el proceso docente educativo debe garantizar el enfrentamiento del estudiante con los problemas que ha de resolver una vez graduado".³⁹ Aspecto que en la Educación Médica cubana se perfecciona cada día más, demostrado con la implementación y desarrollo de la Educación en el Trabajo como forma fundamental de organización de la enseñanza en la formación de los profesionales de la salud.

El profesor ayuda al alumno a aprender determinados contenidos y a su vez a lograr el control progresivo sobre el propio aprendizaje.

Los alumnos son los principales protagonistas del aprendizaje, y desde el punto de vista psicológico es de un gran valor formativo su participación activa y responsable en la evaluación, a medida que los alumnos intervienen en este proceso, profundizan en su propio conocimiento y se convierten en agentes activos y autónomos de su propia formación.⁴²

En el PDE es de vital importancia no solo la relación existente entre los componentes personales, si no la relación lógica entre los componentes no personales o procesales de dicho proceso: objetivos, contenidos, métodos, medios, formas de organización y evaluación.^{6, 20, 37, 40, 41.}

El objetivo es la categoría rectora y la evaluación el mecanismo regulador que puede considerarse como el instrumento de control de la calidad del producto resultante del proceso. Estos componentes deben ser analizados con un enfoque sistémico para evaluar sus interrelaciones e interdependencias.

La evaluación como componente indisoluble de este proceso, parte de los objetivos educacionales como categoría rectora, se interrelaciona dialécticamente con cada uno de los restantes componentes del proceso docente, aspectos que justifican la carencia de sentido de su análisis aislado. Permite valorar la eficiencia real del proceso docente educativo y se convierte en el instrumento eficaz para garantizar su continuo perfeccionamiento y con ello la calidad de la enseñanza.

De lo anterior se impone la necesidad de comprobar a lo largo de toda la carrera se en qué medida preparamos al estudiante para enfrentarse con éxito a su responsabilidad social y si su aprendizaje se corresponde con los objetivos de cada asignatura, disciplina, año de estudio y perfil del egresado.

La evaluación.

Los avances de esta categoría didáctica a lo largo de la historia están encaminados a perfeccionar permanentemente la educación. En la revisión bibliográfica realizada se muestra el interés en la comunidad científica por el referido tema.

Los antecedentes más lejanos de la evaluación están vinculados al examen, y se sitúan en la China Imperial (siglo III a.c.), cuando se realizan varias pruebas de habilidades (manejo del arco, caligrafía) prácticas y académicas para combatir el trato de favor (nepotismo) en la selección de los funcionarios del Estado. Con este mismo objetivo se introdujeron los exámenes en el mundo occidental durante el siglo XIX, para la selección supuestamente más válida y efectiva tanto para entrar en la educación superior como para acceder a determinados puestos de trabajo o ejercer ciertas funciones sociales, políticas, o de otro tipo.⁴³

A finales de siglo XIX y principios del XX existió gran aceptación de las técnicas para la medición del aprendizaje desarrolladas por Edward L. Thondike, quien en 1903 publicó un texto sobre medición del aprendizaje.⁴⁴

En 1942, Ralph Tyler que es considerado por muchos el padre de la evaluación educativa, define la evaluación como el "proceso que permite determinar en qué

grado han sido alcanzados los objetivos educativos propuestos".⁴⁵

Cronbach (1963), propone como alternativa una evaluación que se oriente fundamentalmente a buscar información y comunicarla a quienes han de tomar decisiones sobre la enseñanza, la define como "la recogida y uso de información para tomar decisiones sobre un programa educativo".⁴⁶

Para Scriven citado por Miriam González, la evaluación es la determinación sistemática y objetiva del valor o mérito de algún objeto"⁴⁷

David Stufflebeam (1971) citado por Cerda Gutiérrez, la define como el enjuiciamiento sistemático de la valía o mérito de un objeto.⁴⁸

Lafourcade expone, "La evaluación es una etapa del proceso educacional que tiene que comprobar de modo sistemático, en qué medida se han logrado los objetivos que se han especificado con antelación".⁴⁸

Según Fernández Falcón (1992) la evaluación es una extensión de la investigación educativa que incorpora de forma explícita un componente axiológico (de valores) como parte esencial de interpretaciones o juicios, como un instrumento mediador del propio currículo entre su diseño y puesta en acción.⁴⁹

García Ramos (1996) se entiende por evaluación, en sentido general, aquel conjunto de procesos sistemáticos de recogida, análisis e interpretación de información válida y fiable, que en comparación con una referencia o criterio nos permita llegar a una decisión que favorezca la mejora del objeto evaluado.⁵⁰

Achig (1997), considera que "es un juicio de valor de la misión y los objetivos de una determinada institución, en perspectiva del cambio con amplia participación de sus integrantes, que se sustenta en una rigurosa y sistematizada información, con la finalidad de ayudar a la institución para que comprenda su realidad, reflexione y se proyecte hacia niveles de calidad y pertinencia".⁵¹

Salas Perea (1998), define la evaluación como la actividad que consiste en el análisis e interpretación de los resultados de diferentes instrumentos de medición sobre la actuación de un educando o candidato ante una actividad y/o tarea, en comparación con un patrón o modelo estandarizado, y que se efectúa con la finalidad de tomar una buena decisión.⁶

González Pérez, M (1999) la define la evaluación del aprendizaje como la actividad cuyo objetivo es la valoración del proceso y resultado del aprendizaje de los estudiantes, a los efectos fundamentales de orientar y regular la enseñanza y contribuir al logro de las finalidades de la evaluación.⁵²

Boulmetis y Dutwin (2000), citado por Carballo RM y colaboradores, definen la evaluación como “el proceso sistemático de recopilar y analizar datos con el propósito de determinar si, y en qué grado, se han logrado los objetivos propuestos”.¹⁷

Para Cerda Gutiérrez (2000), la evaluación es la acción de juzgar o inferir juicios sobre cierta información recogida directa o indirectamente de la realidad evaluada, atribuir o negar cualidades al objeto evaluado, o simplemente, medir la eficacia de un método o los resultados de una actividad.⁴⁸

Para María Rosa Milán y Homero Fuentes (2004), la evaluación como eslabón del proceso docente educativo, es consustancial a este, tiene la misma naturaleza de proceso como un todo. Si el proceso docente educativo es considerado como un espacio de construcción de significados, entonces el proceso de evaluación tiene que caracterizarse por ser altamente participativo.³⁸

Segura Castillo (2007), la evaluación, desde una perspectiva ética, persigue ser un procedimiento que respeta la dimensión individual y autónoma de las personas, así como su dimensión colectiva.⁵³

Caraballo R y Rico L (2011) plantean que evaluar es, de manera sucinta, recoger y sistematizar información para tomar decisiones.¹⁷

Según el Reglamento para la organización del proceso docente educativo en los centros de Educación Médica Superior, Resolución Ministerial No 15/1988, se define en el artículo 66: La evaluación del aprendizaje es una parte esencial del proceso de enseñanza-aprendizaje y constituye una vía de retroalimentación para su dirección y para el propio estudiante. Implica el control y la valoración de los conocimientos, habilidades, hábitos, así como de los modos de actuación, que los estudiantes adquieren a través del proceso docente-educativo, al comprobar el grado con que se alcanzan los objetivos propuestos.⁵⁴

De las definiciones anteriores, la autora considera elementos esenciales en la definición de evaluación:

- Determinar el grado en que se alcanzan los objetivos.
- Es un proceso sistemático y continuo.
- Permite emitir un juicio independientemente de lo que se evalúa.
- Brinda información que permite realizar valoraciones para la toma de decisiones.
- Tiene efecto individual y colectivo.
- Es un medio de retroalimentación para profesores y alumnos.
- Un instrumento de investigación didáctica.
- Contiene un componente axiológico.

La complejidad de la evaluación es consecuencia directa de la complejidad del propio proceso de enseñanza-aprendizaje. Éste ha dejado hace tiempo de ser algo que pasaba en las aulas, fuera de la mirada de la sociedad, para convertirse en un hecho casi público, en un hecho social, al que cada vez se le exigen o atribuyen tareas más amplias, hace que la educación no se valore tanto como un proceso de transmisión de conocimientos, sino como un proceso formativo en el que se contempla desde el desarrollo de habilidades básicas hasta la adquisición de los comportamientos sociales y la formación en valores, aspectos de relevante importancia en la Educación Médica Superior.

La autora, después de realizar el análisis de diferentes definiciones de evaluación, asume la referida en la Resolución Ministerial 210 del 2007 del Ministerio de Educación Superior que plantea: La evaluación del aprendizaje es un aspecto consustancial al desarrollo del proceso docente educativo. Tiene como propósito comprobar el grado de cumplimiento de los objetivos formulados en los planes y programas de estudio de la educación, mediante la valoración de los conocimientos y habilidades que adquieren y desarrollan los estudiantes; así como, por la conducta que manifiestan en el proceso docente educativo. Constituye, a su vez, una vía para la retroalimentación y la regulación de dicho proceso.⁴⁰

Las comunidades educativas que, conscientes de su importancia, han adoptado una cultura de evaluación, poseen una herramienta poderosa para identificar y reflexionar respecto a las fortalezas y debilidades de sus prácticas educativas y para tomar decisiones que redunden en el mejoramiento del sistema y, por consiguiente, en el éxito de sus estudiantes.

Si no evaluamos correctamente, si no controlamos las posibilidades de error para poder aceptar o rechazar un resultado, nunca podremos conocer la eficacia de una acción formativa, puesto que al no tener resultados fiables no podremos saber qué niveles de conocimientos tienen los estudiantes ni cuáles son los fallos que se producen en el proceso formativo, se repiten dichos fallos. En resumen, estaríamos envueltos en una ineficacia sistemática.⁵⁵

La Argentina Raquel Salim y colaboradores plantean "nuestro desafío no es revisar y reconstruir las prácticas evaluativas universitarias, sino ubicarlas en un espacio central en el proceso de enseñar y aprender para contribuir a la mejora de la calidad educativa universitaria".⁵⁶

Para la autora Olga de Landaluce³⁹ la evaluación puede ser de Resultado o de Proceso: en el primer caso se encuentran los que consideran verificar al final del proceso los resultados respecto a los objetivos, aunque observen el desenvolvimiento del estudiante durante el proceso no lo recogen, se conforman solo con el resultado final. En la segunda concepción, Evaluación de Proceso, se refiere a los que verifican continuamente sobre la marcha y recogen la información para emitir un juicio de valor.

La considera también como Evaluación cuantitativa y cualitativa: la cuantitativa se refiere a la obtención exclusiva de resultados numéricos, no se da espacio para la reflexión del estudiante. Por su parte la cualitativa, analiza los cambios que se efectúan en el estudiante, en relación con el aprendizaje académico y laboral, así como el desarrollo de su personalidad en el período en estudio.³⁹

Es importante destacar que una preocupación constante de la evaluación cualitativa es el proceso, importa ver cómo se desenvuelve el sujeto durante y a través de todo el proceso docente educativo.

Zabalza en su estudio Competencias Profesionales del Docente Universitario en la Universidad La Salle plantea que evaluar no es conocer algo, ni tener una opinión sobre algo y decirla en alto; es un proceso sistemático que implica como mínimo la recogida de información: para acumular informaciones o datos y disponer del caudal de información suficiente de la realidad a evaluar como para proceder a su evaluación efectiva; la valoración de la información recogida: con la aplicación de los criterios o procedimientos que resulte oportuno se emite un juicio

sobre el valor y pertinencia de los datos disponibles; y la toma de decisiones oportunas. Considera que constituye una de las competencias profesionales del Docente Universitario.⁵⁷

En la educación médica, la evaluación del aprendizaje siempre estuvo condicionada a las costumbres y exigencias sociales. Con el desarrollo de la sociedad e influenciada por diferentes posturas filosóficas, psicológicas y pedagógicas su estudio se dirige, por lo general hacia algunos de sus aspectos como son sus funciones, conceptos, formas y tipos de evaluación.⁶

Estudios realizados, han dado como resultados que la tendencia mundial de la evaluación en la educación médica está dirigida, fundamentalmente hacia:

- La priorización de la evaluación de la competencia y el desempeño profesional como proceso básico para determinar el nivel de desarrollo de los recursos humanos en salud, establecer procesos de certificación y recertificación académica, así como para identificar las necesidades de aprendizaje.
- Que los procesos evaluativos tengan como fin esencial, promover y motivar el desarrollo profesional, en vez del hasta ahora tradicional propósito punitivo.
- La combinación de diferentes métodos, procedimientos y técnicas evaluativas en todos los procesos de análisis, lo que asociado a la integración de una evaluación formativa con la acumulativa o final, permitan el desarrollo de un análisis cada vez más cualitativo de los resultados alcanzados.

Las **funciones** de la evaluación del aprendizaje a nivel mundial ocupan un lugar importante en la literatura pedagógica. Son estructuradas de diferentes formas por determinados autores, aunque expresan ideas parecidas. Con fines didácticos las funciones descritas por varios de ellos resultan las siguientes.⁶

Función de retroalimentación: Está directamente vinculada con la dirección del aprendizaje. En diferentes momentos del proceso docente establecemos cómo se encuentra el aprendizaje de los estudiantes en correspondencia con los objetivos establecidos, esta información retroalimenta al profesor, le señala como está el proceso de asimilación del contenido de la enseñanza y en dependencia de esto

determinará las correcciones necesarias a introducir en la estrategia docente para acercar cada vez más los resultados a las exigencias de los objetivos.

Función instructiva: La evaluación continua contribuye a incrementar, sistematizar, profundizar y generalizar los conocimientos y las habilidades, o sea, la actividad cognoscitiva de los estudiantes; propicia el trabajo independiente y el auto estudio, con lo cual crea las condiciones para la consolidación de los modos de actuación y hacer más efectivo el proceso del aprendizaje.

Función de comprobación y control: La evaluación nos informa sobre el grado en que se cumplen los objetivos de la enseñanza (educativos e instructivos), o sea, nos permite controlar el estado del proceso docente y la efectividad y calidad de la enseñanza y del aprendizaje. Por tanto, la evaluación nos posibilita verificar si los estudiantes han adquirido el nivel de preparación requerido conforme con los objetivos establecidos y a su vez valorar la calidad de los planes y programas de estudio, así como de la estrategia empleada, lo que nos asegura el continuo perfeccionamiento del proceso docente educativo.

Función educativa: La evaluación correctamente organizada y aplicada, donde los estudiantes demuestran los resultados de su esfuerzo ante el profesor y su grupo de estudio, contribuye a que incremente su responsabilidad en el aprendizaje, se plantee mayores exigencias y constituya un motor impulsor, consciente y voluntario en el aumento de la calidad del auto estudio y del trabajo independiente.

No tiene sentido la evaluación, sino como resultante del conjunto de relaciones entre los objetivos, los métodos, el modelo pedagógico, los estudiantes, la sociedad y el docente. Cumple así una función en la regulación y el control del sistema educativo, en la relación de los estudiantes con el conocimiento, de los profesores con los estudiantes, de los estudiantes entre sí, de los docentes y la familia.^{39, 44.}

La autora es del criterio que todas estas funciones están interrelacionadas, tienen en las Universidades un carácter profesionalizador y de acreditación.

La evaluación es un poderoso instrumento para clarificar las metas, conferir sentidos y movilizar a los estudiantes en el logro de los objetivos de la educación.⁵²

Según Salas Perea, para que la evaluación cumpla sus funciones, resulta indispensable la participación del estudiante como sujeto activo del aprendizaje, no solo en los momentos aislados, en que se efectúa un control (lo cual también es muy importante), sino que es necesario concebirla como parte integrante, y a todo lo largo del desarrollo del proceso docente.⁶

Para el estudiante ello supone un autocontrol y autoevaluación a desarrollarse en dos vertientes que se complementan:

- El estudiante compara los resultados que ha obtenido en el control realizado con los que debía haber logrado, y determina sus errores e insuficiencias (control por el resultado). A esto contribuye el análisis que debe realizar el profesor de forma individual o colectiva de los principales problemas que se confrontaron.
- El control del proceso, donde el estudiante precisa en qué momento o fase del proceso se produjeron los errores.

La autora comparte el criterio del citado autor en que para los profesores adquiere importancia el trabajo que en este sentido se realiza en el departamento docente y sobre todo en el colectivo de asignatura o disciplina, donde la participación colectiva del grupo de docentes en la preparación de cada control parcial y final, así como en el análisis de los resultados, asegura la validez y confiabilidad de los mismos; además se mantienen informados de la marcha del proceso, dónde están las dificultades y deficiencias, para la reorientación del trabajo.

Este trabajo de los colectivos de asignatura y disciplina posibilita el continuo perfeccionamiento de la maestría pedagógica de los profesores, pues el estudio y análisis del sistema de evaluación del aprendizaje debe constituir una de las líneas principales a desarrollar mediante las distintas formas y tipos de trabajo metodológico en el colectivo docente que corresponde.

En estos colectivos es donde se precisa el sistema de evaluación a emplear, se determinan las variantes (métodos) posibles a utilizar en relación con las formas y tipos de control a aplicar y los contenidos objeto de control (su nivel de profundidad y de asimilación) de acuerdo con los objetivos a verificar, se confecciona cada instrumento evaluativo con el empleo de los procedimientos y

las técnicas más idóneas para su verificación y definen los criterios para su calificación.

Los conocimientos científicos y tecnológicos, en el mundo contemporáneo en que vivimos, se hacen obsoletos con gran rapidez y no todos los profesionales son conscientes de la necesidad de mantener una búsqueda activa y permanente de la superación profesional que se requiere para poder garantizar la calidad de los procesos educacionales y de la atención médica se brinda a la población.⁵⁸ Implica que los colectivos, departamentos e instituciones deben constituir motores impulsores en tales fines.

El sistema de evaluación se debe diseñar de forma colegiada, con carácter de proyecto, que oriente y regule de modo abierto y democrático, donde todos los implicados tienen el derecho a comprender los argumentos que la sustentan y tomar parte en las decisiones.^{39, 52.}

En la asignatura Metodología de la Investigación y Estadística el sistema de evaluación está formado por evaluaciones frecuentes, parciales y final que como un todo constituyen la evaluación que el estudiante obtiene al concluir la misma.

La autora considera importante destacar además, que este complejo eslabón del proceso docente educativo que involucra al docente, al estudiante y al conocimiento mediado por un determinado contenido escolar, debe desarrollarse a lo largo de todo el proceso de enseñanza aprendizaje, mediante la aplicación de diversos instrumentos que aporten información continua sobre su desarrollo, culmina con evaluaciones finales, y se extiende hasta su formación posgraduada.

El control durante el proceso.

La Resolución 210 de Ministerio de Educación Superior plantea en uno de sus artículos que control del proceso docente educativo es parte del contenido del trabajo metodológico y una de sus funciones. Es el medio fundamental para conocer la calidad de dicho proceso, evaluar sus resultados y dirigirlo hacia el cumplimiento de sus objetivos, comprende, en lo fundamental.⁴⁰

- El control a la planificación y organización del proceso docente educativo en sus distintos niveles organizativos.
- El control a la calidad de la ejecución del proceso docente educativo, que incluye el que ejerce el profesor sobre el aprendizaje de los estudiantes, el

control de la actividad docente, el cumplimiento de los programas de estudio, el aseguramiento material y humano de las actividades docentes y las opiniones de los estudiantes y entidades laborales sobre el desarrollo de las diferentes formas organizativas del proceso docente educativo, entre otros.

- El control a la preparación adquirida por los egresados.

Para evaluar el aprendizaje en la educación superior se realizan diferentes tipos de controles en tres etapas o momentos del proceso: control preliminar; control durante el proceso de enseñanza-aprendizaje y control diferido.^{2, 6, 19, 20, 23.}

El control preliminar es aquel que se realiza con el fin de evaluar el nivel de partida del estudiante. Incluye los exámenes de ingreso, que evalúan la posibilidad de un individuo de ser aceptado en un centro de educación superior o en determinada carrera, porque tienen las condiciones para asimilar con éxito ese tipo de enseñanza, y los exámenes de diagnóstico que se realizan previo al inicio de una asignatura o materia, evalúan el nivel de desarrollo inicial que tiene el estudiante.

El control durante el proceso de enseñanza-aprendizaje se realiza en el transcurso y al culminar una disciplina o asignatura, o al concluir la carrera. Consta de dos tipos de evaluaciones:^{6, 40.}

Formativa, se desarrolla durante el proceso docente, tiene un carácter diagnóstico y nunca se empleará por el profesor para emitir un juicio certificativo, incluye los controles frecuentes y parciales.

Acumulativa o sumativa, se desarrolla al concluir una asignatura, ciclo, carrera y especialidad, tiene un carácter certificativo, integra la evaluación formativa con el ejercicio evaluativo final que se desarrolla.

El control diferido, se realiza algún tiempo después de concluida la enseñanza de una disciplina, asignatura, ciclo o la carrera y tiene como objetivo comprobar la retención de los conocimientos; las habilidades y los modos de actuación adquiridos en ella.⁶

Estas formas de conjunto, caracterizan a la evaluación como un sistema, en correspondencia con su carácter continuo, cualitativo, integrador y basado fundamentalmente en el desempeño del estudiante. La tendencia que debe predominar en el sistema de evaluación es dar el peso fundamental a las

actividades evaluativas frecuentes y parciales, así como evaluaciones finales de carácter integrador.

La autora considera que el sistema de evaluación debe estar concebido de modo que permita la valoración de la efectividad del proceso de formación para actuar en correspondencia a la situación presente, de lo contrario, se convierte en un juicio final de la asignatura, que no representa la realidad del proceso, con errores en su planeación y poco nexo con el modelo del profesional a egresar.

Es de interés del presente estudio el control durante el proceso, específicamente el control final en la asignatura Metodología de la Investigación y Estadística.

Este tipo de control comprueba el logro de sus objetivos más generales y esenciales. Cumple las distintas funciones de la evaluación, excepto la de retroalimentación. Su forma más usual es el examen final, que para la asignatura en estudio se elabora en la Facultad de Ciencias Médicas de Holguín por el metodólogo y el profesor principal, se solicitan propuestas a los profesores que imparten la asignatura en la provincia. Este examen es aprobado por el jefe de departamento.

Para realizar un control con calidad es imprescindible la selección anticipada del instrumento a aplicar. Los instrumentos de evaluación son las herramientas que utiliza el profesor, necesarias para obtener evidencias de los desempeños de los alumnos en el proceso de enseñanza aprendizaje. No son fines en sí mismos, pero constituyen una ayuda para obtener datos e informaciones del estudiante, por ello el profesor debe poner mucha atención en la calidad de éstos ya que un instrumento inadecuado provoca una distorsión de la realidad.⁶

Se pueden definir los instrumentos como los recursos que se emplean para recolectar y registrar información y deben poseer ciertas condiciones para que se garantice la validez, la confiabilidad, la practicidad y otros elementos típicos, de una evaluación de calidad.⁵⁹

La autora María Rosa Alfonso considera que un instrumento de evaluación es todo recurso concreto que el docente utiliza para reunir información sobre el proceso de aprendizaje de los alumnos y sus resultados.⁶⁰

Planeación del instrumento.

Es de primordial importancia que la planificación de la Evaluación del aprendizaje sea un proceso compartido a nivel institucional entre todos los profesores.⁵³

Según autores como Salas Perea⁶ y otros^{2, 19, 20, 23.} para la confección de un instrumento evaluativo se deben definir 4 etapas:

1. Definir qué debe ser evaluado.
2. Seleccionar los métodos y procedimientos a emplear.
3. Cómo calificar un examen.
4. Cómo establecer un patrón de aprobado/desaprobado.

La primera etapa; definir qué debe ser evaluado; constituye la base de la validez de contenido del instrumento que se confecciona. Es necesario identificar los problemas que los educandos deben ser capaces de enfrentar y/o resolver; definir para cada problema las tareas en las que se espera sea competente y preparar una tabla de especificaciones o de contenidos para determinar la selección de los problemas y tareas que deben ser incluidos en el instrumento evaluativo.⁶

La segunda etapa; seleccionar los métodos y procedimientos a emplear; se realizará sobre la base de los objetivos y contenidos a evaluar. En su selección debemos tener presente:

- Emplear con preferencia aquellos que pongan a los educandos en situaciones reales o lo más parecido posible.
- La tarea clínica a evaluar prácticamente indica los métodos, procedimientos y técnicas a emplear.
- Precisar cómo complementar la tarea práctica con los conocimientos que debemos evaluar y seleccionar en consecuencia con los procedimientos y técnicas a emplear.
- Valorar las restricciones que pueden condicionar su selección como preparación del claustro, tiempo disponible para el examen, cantidad de estudiantes a examinar; recursos existentes, entre otros.

En la tercera etapa; cómo calificar un examen; se emplean determinadas técnicas que deben estar adecuadamente diseñadas con suficiente anterioridad y validadas inicialmente por consenso de expertos y posteriormente por el análisis matemático de sus resultados. Entre las más empleadas tenemos listas de

comprobación y escalas de calificación; claves de calificación de respuestas, tablas de decisiones entre otras.

La calificación y la obtención de una nota favorece la evaluación formativa en el aprendizaje de los estudiantes cuando dejan de ocupar el lugar central y la atención de los implicados, se desplaza al proceso evaluativo y sus diversas funciones, a valorizar el conocimiento y no la nota.⁶

La cuarta etapa; cómo establecer un patrón de aprobado/desaprobado; es un proceso complejo a desarrollar por el colectivo docente, con el objetivo de definir los elementos que determinan la calificación de aprobado en un examen y el orden de los elementos que caracterizan a cada escala de calificación utilizada.

Se han desarrollado diferentes enfoques a partir de la propuesta de Livingston y Ziesky en 1982, citado por otros autores que podemos resumir en dos grandes grupos:⁶

Enfoque normativo: Establece un patrón relativo que depende de la competencia mostrada por cada examinado. La calificación del aprobado se establece con posterioridad al examen. La escala puede ser numérica o alfabética.

Enfoque criteriado: Parte de un patrón absoluto, que se establece con anterioridad a la realización del examen. La selección de la calificación del aprobado es arbitraria e independiente de los resultados que obtengan los estudiantes.

En la medida que podamos construir buenos instrumentos de evaluación, podemos medir, con todo el rigor posible, las competencias de los profesionales que se desempeñan en el sistema nacional de salud. Esta es la capacidad de utilizar los conocimientos, las habilidades y las cualidades interpersonales para resolver los problemas cotidianos encontrados en su práctica profesional.⁶

Existen varios sistemas para categorizar las operaciones cognitivas que son usadas en el desarrollo de las especificaciones del instrumento.

Según plantea Sala Perea,⁶ internacionalmente la más conocida es la taxonomía de Bloom (1956) que contiene las categorías: conocimiento, comprensión, aplicación, análisis, síntesis y evaluación. Thorndike y Hagen (1977), proponen categorías como: recordar, reconocer, identificar, definir, aplicar, analizar, sintetizar, generalizar, predecir o evaluar.

En Cuba se emplean cuatro niveles de asimilación: familiarizarse, reproducir, producir y crear.⁶

El primer nivel o familiarizarse, implica que el educando tiene una idea de conjunto o noción del contenido que le es informado aunque no pueda reproducirlo ni utilizarlo, no es capaz de resolver ninguna situación que se le presente.

El segundo nivel o reproducir, implica que conoce el contenido o la situación problema planteada y es capaz de repetirlo.

En el tercer nivel o producir, el estudiante está preparado para aplicar el contenido ya conocido ante una situación o problema nuevo; o sea está preparado para saber emplear el conocimiento o habilidad aprendida.

El cuarto nivel o crear, supone la capacidad del estudiante para enfrentar y resolver situaciones nuevas para las que no dispone de todos los elementos.

Al confeccionar instrumentos evaluativos es importante reconocer que en una pregunta o problema pueden incidir diferentes niveles de procesos cognitivos y se debe lograr proponer tareas que movilicen potencialmente competencias.

La aplicación de las diferentes formas de evaluación se puede realizar mediante múltiples métodos, procedimientos y técnicas. Los principales métodos empleados en la confección de los instrumentos son: teóricos, prácticos y teórico-prácticos.⁶

Los métodos teóricos se emplean para evaluar conocimientos y habilidades intelectuales, pueden ser orales y escritos. Los métodos prácticos se emplean para evaluar las habilidades manuales o destrezas y modos de actuación. Los orales se pueden desarrollar mediante los procedimientos siguientes: el examen oral, la entrevista y la consulta docente educativa fundamentalmente. Los escritos pueden dividirse en dos grandes familias:^{2, 6, 19, 20, 59, 60.}

- Los exámenes de desarrollo, de respuesta libre, tradicionales o de ensayo.
- Los exámenes de respuesta estructurada o test objetivos.

Producto de la combinación de estas dos variantes se puede definir otra variante llamada examen mixto.

Esta clasificación de los exámenes se basa en la categoría de las preguntas o ítems que pueden ser de respuesta libre o ensayo y de respuesta estructurada o test objetivo.

El primer tipo es aquel en el cual el sujeto aporta su propia respuesta, por ejemplo la pregunta de ensayo que requiere de una respuesta extensa; la pregunta de respuesta corta que no necesita más que una o dos oraciones para responderla; y la de completamiento que sólo requiere de una palabra o frase para responderse, entre otras.

El segundo tipo es aquel en el cual el sujeto elige su respuesta dentro de varias respuestas proporcionadas por el elaborador del examen, tales como el ítem de respuesta alternativa (verdadero o falso); el ítem de selección múltiple (complemento simple o agrupado) y el ítem de igualamiento o pareación, entre otros.

Es reconocido por diferentes investigadores y la autora de este estudio que las pruebas de desarrollo, permiten conocer la profundidad en el dominio de una temática; pero tienen la dificultad que es escasa la muestra del contenido a evaluar y la subjetividad del evaluador influye en gran medida en la puntuación otorgada al sujeto, requiere de mucho tiempo para responderla y no permiten la estandarización de los resultados al ser difícil elaborar claves de calificación.

La comunidad investigativa considera que todos los tipos de ítems, poseen sus ventajas y desventajas.^{2, 5, 18- 20, 54- 55.}

Ventajas de las preguntas tipo ensayo:

- Posibilita que el educando exponga sus conocimientos, ordene sus ideas y demuestre su capacidad para expresarse en su idioma.
- Requiere poco tiempo para redactar las preguntas.
- Mide la capacidad de resolver problemas novedosos.
- Mide la capacidad de organizar, integrar o sintetizar.
- No se puede adivinar la respuesta sin conocerla.
- Potencia la capacidad de expresión, organización, originalidad, etc.

Entre sus desventajas se encuentra que:

- Limita sensiblemente el número de áreas a explorar.
- La objetividad es difícil de conseguir.
- Proporciona escasa retroalimentación al educando y al claustro.

- Su calificación consume demasiado tiempo, en ocasiones influye la mala letra, es subjetiva y sus resultados son poco confiables.

Según la literatura consultada existen indicaciones generales para su uso entre las que se encuentran que el grupo de estudiantes sea pequeño; los objetivos del control contemplen evaluar las habilidades de autoexpresión escrita; que los examinadores se sientan satisfechos con evaluar a través de pocas preguntas una muestra de los contenidos que deben ser evaluados; que los examinadores tengan experiencia y aplicarla cuando es más importante evaluar la libre expresión del estudiante que la objetividad y confiabilidad de una calificación

Por otra parte, las preguntas de ensayo de respuestas cortas o restringidas, reducen las desventajas de las preguntas de tipo ensayo, se incrementa la muestra de contenido a evaluar y se limitan las preguntas sólo a aquellas que conlleven respuestas libres, pero cortas. Pueden tener forma de pregunta de respuesta breve o de un enunciado con espacios en blanco para llenar o completar.

En general, para designar una pregunta como de respuesta restringida debe poder contestarse con una sola palabra, una fecha, un número, una frase o a lo sumo con no más de dos o tres oraciones.

Entre sus ventajas se encuentran: las preguntas son fáciles de confeccionar, se incrementan las áreas de materias a examinar al aumentar el número de preguntas, es más fácil para los examinadores ponerse de acuerdo en las posibles respuestas a aceptar como correctas y en cómo calificarlas, la objetividad y confiabilidad de las calificaciones se incrementa considerablemente.

Como desventajas se reconocen: la amplitud de las áreas / materias a examinar es aún limitada, la calificación aún consume mucho tiempo y persiste el juicio subjetivo alrededor del límite de lo correcto/incorrecto de las respuestas emitidas por los educandos, las respuestas solicitadas son frecuentemente reproductivas más que de razonamiento y discusión.

En sentido general, para el uso de preguntas de ensayo es importante tener en cuenta reservar las preguntas para comprobar resultados que no se puedan comprobar mejor mediante pruebas objetivas, formularlas con mucho cuidado para que centren la atención del alumno en la tarea deseada, la que debe estar bien definida.⁵⁹

Otra variante de preguntas de ensayo lo constituye la de ensayo modificado cuyo formato se basa en un problema o situación donde los elementos son presentados en etapas sucesivas. Se utilizan como contenidos fundamentales problemas o situaciones de salud.

A partir de la presentación de la situación problema inicial se formulan una serie sucesiva de etapas / problemas que constan de una batería de preguntas de ensayo tradicional o de ensayo de respuestas breves acerca del desarrollo de la situación, donde las respuestas en cada etapa son relativamente libres, ya que no son sugeridas al evaluado.

Permiten explorar, entre otros aspectos, la identificación de los factores claves que influyen en la toma de una decisión y sus actitudes correspondientes, en cuanto a su conducta ante el trabajo profesional diario. Las preguntas no contienen aspectos explícitamente señalados. Se espera que los estudiantes enfrenten el problema de modo que apliquen lo aprendido.⁶

Entre las ventajas referidas en la literatura se encuentran: el empleo de descripciones clínicas incrementa la validez del examen para medir competencias intelectuales; pueden evaluarse las conductas a realizar ante un problema no diferenciado o general; los profesores pueden identificar las áreas o materias débiles en cada estudiante y grupo; la conducta a seguir puede ser evaluada a través de diferentes fases, en cuanto a diagnóstico, investigaciones, tratamiento y consejos al paciente y su familia y con un adecuado diseño se puede incrementar la validez del examen.

Presentan también desventajas entre las que se encuentran: cada caso problema requiere tiempo de preparación y realización del examen; cada caso problema representa un examen en sí mismo, y la información que se brinda es limitada, lo que pudiera ser insuficiente y limitarle su interpretación y respuesta; su calificación requiere tiempo y juicio profesional. Los examinadores deben sopesar cada pregunta y ponerse de acuerdo, previamente, ante las posibles respuestas y valorar cuáles serán aceptadas como correctas.

Las preguntas de test objetivo son preferidas por muchos evaluadores por la facilidad, la rapidez y la objetividad de su calificación.

Este tipo de preguntas es útil para examinar un grupo grande de estudiantes, un amplio contenido muestral a evaluar, necesidad de obtener calificaciones muy confiables, evaluación imparcial, justa y libre del “efecto de halo” con necesidad de divulgar los resultados con prontitud.

Presentan ventajas como:

- Asegura objetividad y, en general, propicia el aumento de la confiabilidad y validez del instrumento.
- Incrementa la gama y variedad de los aspectos a evaluar, en poco tiempo y con mínima supervisión.
- Posibilita la facilidad y rapidez de su calificación, lo que permite divulgar con prontitud las calificaciones obtenidas.
- Suministra retroalimentación, tanto al profesor como a los estudiantes, evidencian los puntos débiles.
- El análisis estadístico es relativamente sencillo y suministra información confiable.
- Pueden almacenarse bancos de preguntas.

Entre sus desventajas se encuentran:

- Su preparación exige mucho tiempo, competencia y práctica con el fin de evitar preguntas ambiguas y arbitrarias.
- Proporciona a los educandos informaciones e indicaciones que no dispondrían en una situación real.
- Resulta costoso al examinar pocos estudiantes.
- Implica adoptar determinadas condiciones para tener en cuenta las respuestas correctas que pudieran ser adivinadas al azar.
- Exige el entrenamiento previo de los educandos.

Algunos autores plantean que las pruebas de respuesta abierta constituyen una modalidad muy popular y un instrumento clásico de evaluación del aprendizaje. Se consideran procedimientos mixtos de evaluación al ser aplicadas conjuntamente con las pruebas objetivas.^{6, 60.}

La autora considera que todos los tipos de ítems o preguntas poseen su utilidad a la hora de elaborar un instrumento de evaluación, su estudio constituye interés permanente de los investigadores y de los claustros de profesores en aras de buscar las mejores soluciones para elevar la calidad de los instrumentos que se

aplican, todos poseen ventajas y desventajas, el éxito de su uso depende del conocimiento que se tenga de cada uno de ellos, para así poder elevar sus ventajas y reducir sus desventajas.

En general, todos los instrumentos evaluativos deben reunir las siguientes características: sus contenidos deben estar directamente relacionados con los objetivos educacionales, ser realista y práctico, tener validez y confiabilidad, ser objetivo y factible de realizar, comprender las cantidades importantes y útiles, ser completo, pero lo más breve posible, ser preciso, pero claro en su redacción.⁶

El examen final de la asignatura Metodología de la Investigación y Estadística se realiza en la Facultad de Ciencias Médicas de Holguín de forma escrita, la calificación se realiza con un enfoque criteriado, se clasifican en aprobados y desaprobados con una escala de 2 a 5, resultan desaprobados los que obtienen calificación de 2.

Calidad del instrumento.

Al concluir la confección de un instrumento evaluativo es necesario determinar en qué medida la información obtenida refleja el nivel de competencia del examinado.

Guilbert citado por Salas de Perea⁶ recomienda tener en cuenta al evaluar la calidad de los instrumentos evaluativos conceptos como:

- **Objetividad:** grado de concordancia entre los juicios emitidos por evaluadores independientes y competentes sobre lo que constituye una "buena" respuesta para cada una de las preguntas de un instrumento evaluativo.
- **Pertinencia:** es el grado de respeto de los criterios establecidos en la selección de las preguntas para que sean conformes con los fines y objetivos de la evaluación.
- **Equilibrio:** grado de concordancia entre la proporción de preguntas reservadas a cada uno de los objetivos y la proporción ideal para un instrumento de este tipo.
- **Equidad:** grado de concordancia entre las preguntas planteadas en el instrumento evaluativo y el contenido que se pretende evaluar.
- **Discriminación:** cualidad de cada pregunta del instrumento evaluativo que permite distinguir los estudiantes de altos y bajos rendimientos.

- **Eficacia:** cualidad de un instrumento evaluativo que permite el mayor número posible de respuestas independientes por unidad de tiempo.

La correcta aplicación de un sistema de evaluación del aprendizaje, que tiene un carácter cualitativo integrador, se basa en un conjunto de condiciones que constituyen las premisas básicas que se deben satisfacer, las cuales se encuentran estrechamente relacionadas entre sí y cuya inobservancia impide que se cumplan efectivamente sus funciones. Estas premisas se estructuran en principios que son: objetividad, sistematización y continuidad, carácter sistémico, y exigencias como son: la validez y la confiabilidad.⁶

Para que una evaluación sea objetiva es imprescindible, que parta de los objetivos definidos, determine el dominio del contenido de la enseñanza y las cualidades que lo caracterizan, de acuerdo con el nivel de profundidad precisado y su correspondencia con el nivel de asimilación exigido y que a su vez esta valoración se realice mediante la forma y tipo de evaluación que asegure la validez del control.^{6, 52.}

La sistematización y continuidad de la evaluación está dada porque el número de muestras (controles) del universo de las actividades docentes que realice el estudiante y la regularidad de su realización, posibilite la generalización del criterio del nivel de aprendizaje alcanzado. Es necesario caracterizar el aprovechamiento del estudiante con un criterio dinámico, evolutivo y activo.

La frecuencia debe ser tal que permita ayudar oportunamente al estudiante y orientarlo al fin deseado y retroalimente al profesor, también debe estar en dependencia del éxito que se logre en el estudiante durante el desarrollo del proceso docente y de la necesidad que tenga de ser sometido al control.

Su carácter sistémico está dado en la necesidad de concebirlo como un componente del proceso de enseñanza-aprendizaje y en estrecha interrelación e interdependencia con los restantes componentes del mismo: objetivo, contenido, método, formas y medios de enseñanza, así como en el papel activo que tienen que desarrollar tanto el profesor como el estudiante para su efectividad.

Según Salas de Perea⁶ un instrumento evaluativo tiene calidad en la medida en que es válido y confiable y sus resultados son generalizables.

La validez, es la correspondencia entre lo que se pretende verificar por el instrumento, procedimiento o método y lo que realmente se mide o evalúa. Abarca tanto a las exigencias propias del control como a la determinación de los índices valorativos que permiten la calificación. Mientras más teórico, reproductivo y alejado de la práctica esté un control mucho menor será su validez. Requiere del criterio de expertos. Usualmente se clasifica en cuatro grupos: validez de contenido, validez de criterio, validez de construcción y validez funcional.⁶

La validez de contenido o conceptual está en correspondencia con los conocimientos y las habilidades (tanto cualitativa como cuantitativamente) que evalúa o mide el instrumento de control sean los suficientes para definir el grado de aprovechamiento del estudiante, sobre la base de los objetivos propuestos.

En un estudio de validez de criterio, se comparan los resultados de la aplicación del instrumento evaluativo con otros criterios externos que se le llama "variable-criterio". Se dice que un instrumento evaluativo tiene validez de criterio si, a partir de los resultados obtenidos, se pueden hacer inferencias válidas sobre una variable que no se puede medir directamente por el instrumento evaluativo.

Los estudios de validez de criterio se clasifican en: validez predictiva que indica la efectividad del instrumento evaluativo en un resultado futuro. Para esto se comparan las calificaciones del instrumento evaluativo con una medida directa de la ejecución posterior de los sujetos llamado criterio y la validez concurrente se refiere a la interrelación del instrumento evaluativo y un criterio de medición efectuado en el momento en que se aplica el instrumento evaluativo.

Los estudios de validez de construcción de un instrumento evaluativo se realizan cuando no se dispone de una medida directa del atributo a medir, de tal manera que hay que sustituirlo por un modelo hipotético.

La validez funcional está en correspondencia con la selección del tipo de actividad (acciones, tareas, problemas) que se deben controlar en el estudiante, para evaluar sus modos de actuación de acuerdo con los objetivos propuestos. Ambos tipos de validez se interrelacionan y complementan.⁶

La autora Miriam González Pérez⁵², hace referencia a la validez ecológica que acerque la evaluación a las condiciones naturales del proceso de aprendizaje y su contenido.

La confiabilidad es la estabilidad en los resultados de un instrumento evaluativo, ya sea al repetirlo, o al ser calificado por distintos profesores. La validez y la confiabilidad están estrechamente relacionadas. Un control que cumpla con las exigencias de la validez, tiene un alto grado de probabilidad de ser confiable; sin embargo, no necesariamente ocurre así a la inversa.^{6, 61.}

La generalizabilidad consiste en estudiar los componentes del error y determinar la contribución de cada uno de ellos en el error total calculado.

La autora considera importante tener presente que la construcción de un instrumento evaluativo absolutamente seguro es prácticamente imposible en la realidad académica, ya que los instrumentos nunca pueden abarcar todo el conocimiento que se pretende medir.

La confiabilidad es un concepto estrictamente estadístico que se expresa por un coeficiente de confiabilidad o a través del error típico de las mediciones efectuadas.

En la literatura especializada se reportan varios procedimientos para cuantificar la confiabilidad de los instrumentos evaluativos.^{6,61.} Entre los que se encuentran:

- El coeficiente θ .
- El coeficiente ω .
- El coeficiente de Kuder-Richarson.
- El coeficiente α de Cronbach.

El coeficiente Alfa(α) de Cronbach desarrollado por J. L. Cronbach, es el más utilizado en el mundo y según criterio de varios autores es el mejor para estimar el error de muestreo, proporciona una medida efectiva de la consistencia interna del instrumento evaluativo, se puede emplear ante diferentes formatos de preguntas.^{2, 6, 61.}

Entre sus ventajas se encuentra la posibilidad de evaluar cuanto mejoraría (o empeoraría) la fiabilidad de la prueba si se excluyera un determinado ítem o pregunta.⁶³

Requiere una sola administración del instrumento de medición y produce valores que oscilan entre 0 y 1. Un coeficiente de 0 significa nula confiabilidad y 1 representa un máximo de confiabilidad (confiabilidad total).

Autores seguidores de Anastasi y de Weiner aceptan como adecuados en exámenes escritos basados en preguntas tipo test objetivo, los coeficientes que oscilan entre 0.80 y 0.85.⁶

El colegio de Médicos de Familia en Canadá en sus exámenes de certificación de especialistas reporta para los exámenes escritos basados en respuestas cortas entre 0.69 y 0.71, en los exámenes de Simulación Oral en el Consultorio Médico entre 0.54 y 0.63, en los exámenes clínicos objetivamente estructurados (OSCE) con pacientes estandarizados y preguntas de ensayo de respuestas cortas los coeficientes propuestos se encuentran entre 0.71 y 0.77.⁶

Para el examen que se analiza, la autora asume el valor del coeficiente Alfa de Cronbach que se propone en el proyecto de las Indicaciones Metodológicas para el Planeamiento, Diseño y Control de la Calidad de los Exámenes Escritos para la Evaluación de la Competencia y el Desempeño Profesional en Cuba, que considera como confiabilidad aceptable un valor de 0,6.⁶²

Otros indicadores que miden la calidad de un examen escrito son el nivel de dificultad y el poder de discriminación.^{2, 12, 14-16.}

Como plantea Wood (1960), citado por Backhoff¹⁵ el nivel de dificultad de un ítem se entiende como la proporción de personas que responden correctamente una pregunta de una prueba. Entre mayor sea esta proporción, menor será su dificultad, se trata de una relación inversa. A ésta proporción se le denota con la letra p.

El manual del Examen de Habilidades y Conocimientos Básicos (EXHCOBA) de la universidad de Baja California en México considera el nivel medio de dificultad del examen entre 0.5 y 0.6.¹⁵

Córica y colaboradores plantean que un examen que posea dificultad media, es superior, en confiabilidad de resultados, a aquellos exámenes que tengan mucha dificultad o los que resultan muy fáciles, pues dan una mayor información sobre los estudiantes que tienen dominio del contenido evaluado de aquellos que no lo tienen.⁶⁴

La autora de este trabajo comparte el criterio de Enrique Santiesteban Leyva cuando expresa "El Índice de dificultad constituye un parámetro de fácil manejo, que no requiere de cálculos complejos, de la búsqueda de datos adicionales, ni de

procedimientos engorrosos, su utilización resulta recomendable en el análisis de los resultados de los exámenes".²³

El poder de discriminación del examen no es más que la cualidad de cada pregunta del instrumento que permite distinguir los estudiantes con altos y bajos rendimientos.²

Si la prueba y un ítem miden la misma habilidad o competencia, se debe esperar que quien tuvo una puntuación alta en todo el test deberá tener altas probabilidades de contestar correctamente el ítem. También se debe esperar lo contrario, es decir, que quien tuvo bajas puntuaciones en el test, deberá tener pocas probabilidades de contestar correctamente el reactivo. Así, un buen ítem debe discriminar entre aquellos que obtuvieron buenas calificaciones en la prueba y aquellos que obtuvieron bajas calificaciones.¹⁵

Las dos formas más utilizadas para determinar el poder discriminativo de un ítem son el índice de discriminación (D) y el coeficiente de discriminación.¹⁵

Entre más alto es el índice de discriminación, el reactivo diferenciará mejor a las personas con altas y bajas calificaciones. El coeficiente de discriminación se mide a través de dos formas: el coeficiente de correlación biserial y el coeficiente de correlación del punto biserial.

El coeficiente de discriminación ofrece ventajas respecto al índice de discriminación, el primer método toma en cuenta todos y cada uno de los estudiantes evaluados, mientras que el segundo solo toma 54 % de ellos 27 % más más alto y 27% más bajo.¹⁵

El coeficiente de correlación biserial (r_{bis}) se calcula para determinar el grado en que las competencias que mide el test también las mide el reactivo.

El coeficiente de correlación del punto biserial (r_{pbis}) se utiliza para saber si las personas "adecuadas" son las que obtienen las respuestas correctas, qué tanto poder predictivo tiene el reactivo, y cómo puede contribuir a las predicciones.

Henrysson (1971) citado por Backhoff¹⁵ sugiere que el r_{pbis} nos dice más sobre la validez predictiva del test que el coeficiente de correlación biserial, ya que éste tiende a favorecer los reactivos de dificultad media. También se sugiere que el

rpbis es una medida que combina la relación entre el criterio del reactivo y el nivel de dificultad.

Estudio realizado por María Guadalupe Sánchez González¹⁶ en la Facultad de Veterinaria y Zootecnia en México revela la superioridad del coeficiente de correlación del punto biserial frente al índice de discriminación.

A partir de lo expuesto la autora asume en este estudio el uso de los indicadores de calidad índice de dificultad, coeficiente de correlación punto biserial y coeficiente alfa de Cronbach que permitirán realizar la caracterización de la calidad del examen final de la asignatura Metodología de la Investigación y Estadística.

Objetivo general.

Caracterizar la calidad del examen final de la asignatura Metodología de la Investigación y Estadística aplicado a estudiantes de 2do año de la carrera de Medicina, curso 2010 - 2011.

Objetivos específicos.

- 1.- Describir la estructura del examen final aplicado en la asignatura Metodología de la investigación y Estadística.
- 2.- Analizar la correspondencia entre el fondo de tiempo por temas y la cantidad de ítems explorados en el examen.
- 3.- Valorar el comportamiento del índice de dificultad, coeficiente de discriminación del punto biserial y coeficiente alfa de Cronbach en el instrumento aplicado.

Diseño Metodológico.

Se realizó una investigación en el campo de la evaluación del aprendizaje, cuyo objeto fue el examen final de la asignatura Metodología y Estadística en la carrera de Medicina, con el objetivo de caracterizar el instrumento que se aplicó a los estudiantes de segundo año en el curso 2010 - 2011.

El instrumento de análisis fue confeccionado por expertos de la Sede Central de la Facultad de Ciencias Médicas "Mariana Grajales Coello" y se aplicó en el primer semestre del curso escolar 2010-2011.

El Universo estuvo constituido por el total de exámenes finales ordinarios de Metodología y Estadística de los estudiantes de segundo año de la carrera de Medicina de la provincia Holguín en el curso 2010-2011.

Se seleccionó una muestra intencionada de 276 exámenes de los estudiantes presentados a esta convocatoria en el municipio Holguín.

Se utilizaron métodos teóricos, empíricos y procedimientos estadísticos como:

Métodos Teóricos:

Análisis documental: para la revisión del programa y otros documentos normativos de la asignatura y la carrera de Medicina, bibliografía relacionada con el proceso docente educativo, la evaluación del aprendizaje, calidad de los instrumentos de evaluación y otras de interés.

Histórico lógico: para relacionar diferentes momentos y concepciones vinculadas con el desarrollo histórico de la disciplina, de la evaluación del aprendizaje e instrumentos de evaluación en los que se sustentó el marco teórico.

Análisis síntesis: en el procesamiento de la información teórica y empírica, al realizar el análisis de los exámenes, la elaboración de los resultados y conclusiones sobre la calidad del examen aplicado.

Inducción deducción: para arribar a generalizaciones acerca de las potencialidades y limitaciones que presentó el instrumento de evaluación aplicado.

Métodos Empíricos:

Entrevista a expertos: para caracterizar el examen según su formato, los

contenidos que evalúan, el criterio de respuestas correctas en los ítems de tipo ensayo y la clasificación de la escala de dificultad de los ítems esperados. Para ello se seleccionaron 6 profesores de la sede central con más de 10 años de experiencia docente y más de 5 años de trabajo en ciencias médicas, con cargos de dirección y profesores del departamento de Informática Médica con categoría docente de asistente o auxiliar (Anexo 1).

Procedimientos estadísticos:

Para el análisis cuantitativo de los datos que aportó el instrumento evaluativo aplicado y su procesamiento en el cálculo de los diferentes indicadores.

- Descripción del examen aplicado.

Para describir el examen en estudio se tuvo en cuenta el criterio emitido por los expertos y elementos de la guía elaborada por Moreno Montañez² en su investigación:

- Tipo de examen.
- Cantidad de preguntas del examen.
- Cantidad de ítem por preguntas.
- Tipos de ítems que contiene el instrumento según clasificación de Salas Perea.
- Contenido que aborda cada pregunta.

En una hoja de cálculo de Microsoft Excel se tabularon los resultados del examen objeto de estudio, se incluyeron los resultados de cada ítem del examen categorizados en aprobado (1) y suspenso (0), la calificación de cada pregunta y la nota final del examen, para su posterior procesamiento estadístico.

- Correspondencia del fondo de tiempo y los ítems explorados.

Se analizó el tiempo de cada tema en el programa de la asignatura, a partir de esta información se determinó el por ciento que representa cada uno respecto al total de horas y la cantidad de ítems explorados en cada temática con su respectivo por ciento respecto al total de incisos. Para corroborar esta relación se aplicó el Test de Correlación de Spearman.

Se realizó la valoración de los siguientes indicadores de calidad:

- Índice de dificultad.

Se calculó por ítems, temas, preguntas y examen.

Para calcular la dificultad del ítem, se dividió el número de personas que contestó correctamente el ítem entre el número total de personas que contestaron el ítem. Usualmente, a esta proporción se le denota con una p y se utilizó la siguiente fórmula:^{6, 18.}

$$P = \frac{A}{N}$$

Donde p = Índice de dificultad.

A = Número de aciertos.

N = Número de aciertos más número de errores.

Definición de la escala de clasificación de la dificultad.

Con el apoyo de los expertos se elaboró la escala de clasificación del nivel de dificultad de los ítems del examen, se definieron los intervalos según la proporción de estudiantes aprobados en ellos para determinar la clasificación de los incisos en correspondencia con la dificultad real presentada en el instrumento aplicado.

Se determinó que los ítems que tuvieran una proporción de aprobados:

- Menor de 0,40 se consideran difíciles.
- Entre 0,40 y 0,50 se consideran medianamente difíciles.
- Entre 0,51 y 0,80 se consideran de dificultad media.
- Entre 0,81 y 0,90 se consideran medianamente fáciles.
- Entre 0,91 y 1,0 se consideran fáciles.

Escala de clasificación de la dificultad esperada en los ítems:

Es considerado por la literatura que el examen debe tener.⁶

- 5 % para ítems fáciles.
- 20 % para ítems medianamente fáciles.
- 50 % para ítems de dificultad media.
- 20 % para ítems medianamente difíciles.
- 5 % para ítems difíciles.

Se realizó la distribución de los ítems del examen según el Índice de dificultad esperado y observado con su respectivo por ciento. El cálculo del coeficiente de correlación de Spermán se utilizó para corroborar esta relación.

Para calcular el **índice de dificultad de las temáticas** se sumó el total de estudiantes que respondieron correctamente cada uno de los ítems de la temática y se dividió entre el número de ítems que exploró ese tema, se obtuvo la media de respuestas correctas para el tema. A continuación ese resultado se dividió por el total de examinados, como se representa en la siguiente fórmula:

$$D_t = \frac{\sum (A_{t1} + A_{t2} + \dots + A_{tn})}{\frac{n_t}{N_t}}$$

Donde:

D_t = Índice de dificultad de la temática.

A_t = Número de respuestas correctas de los ítems de la temática.

n_t = Número de ítems de la temática.

N_t = Número total de presentados en el examen.

Para calcular el **índice de dificultad de las preguntas** se sumó el total de estudiantes que respondieron correctamente cada uno de los ítems que corresponden a la pregunta y se dividió entre el número de ítems que exploró la pregunta, se obtuvo la media de respuestas correctas para esa pregunta y se dividió por el total de examinados. Se utilizó la fórmula anterior aplicada a la pregunta como se muestra a continuación:

$$D_p = \frac{\sum (A_{p1} + A_{p2} + \dots + A_{pp})}{\frac{n_t}{N_t}}$$

Donde:

D_p = Índice de dificultad de la pregunta.

A_p = Número de respuestas correctas de los ítems de la pregunta.

n_p = Número de ítems de la pregunta.

N_t = Número total de presentados en el examen.

Con la escala de clasificación de la dificultad por ítems, se evaluó la dificultad por temática y por pregunta.

- **Coefficiente de discriminación del punto biserial (r_{pbis}).**

Se calculó para determinar para determinar la efectividad discriminativa de un reactivo. Permite determinar si los estudiantes “adecuados” son los que obtienen las respuestas correctas, qué tanto poder predictivo tiene el reactivo, y cómo puede contribuir a las predicciones.

Se utilizó la siguiente fórmula:

$$r_{pbis} = \frac{\overline{X}_1 - \overline{X}_0}{S_x} * \sqrt{\frac{n_1 \cdot n_0}{n(n-1)}}$$

Donde:

\overline{X}_1 = Media de las puntuaciones totales de aquellos que respondieron correctamente el ítem.

\overline{X}_0 = Media de las puntuaciones totales de aquellos que respondieron incorrectamente el ítem.

S_x = Desviación estándar de las puntuaciones totales.

n_1 = Número de casos que respondieron correctamente el ítem.

n_0 = Número de casos que respondieron incorrectamente el ítem.

$n = n_1 + n_0$

Según estándares internacionales la distribución de los resultados del coeficiente discriminación punto de correlación biserial (r_{pbis}) es:¹⁵

- < 0 discriminan negativamente.
- 0 - 0.14 discriminan pobremente.
- 0.15 - 0.25 discriminan regular.
- 0.26 - 0.35 buen poder discriminativo.
- > 0.35 excelente poder de discriminación.

El r_{pbis} se calculó por ítems y preguntas.

- **Coefficiente α de Cronbach.**

Se utilizó para el análisis de la confiabilidad, es un indicador que permite determinar la consistencia interna del examen.

El alfa de Cronbach toma un valor positivo. Para el examen en estudio se consideró adecuado valores de 0,60 según lo planteado en el proyecto de las Indicaciones Metodológicas para el Planeamiento, Diseño y Control de la Calidad de los Exámenes Escritos para la Evaluación de la Competencia y el Desempeño Profesional.⁶⁴

Se calculó para el examen y al eliminar una pregunta para determinar la influencia de cada una en la consistencia interna del instrumento aplicado.

En los requerimientos éticos de la investigación se consideró la voluntariedad de participación de los expertos implicados.

Finalmente se contrastaron las informaciones obtenidas de la revisión documental, los métodos empíricos y estadísticos, se procedió a la triangulación de los resultados de forma resumida y lógica.

La edición de la tesis se realizó mediante una computadora Pentium 4 con ambiente Windows Xp Profesional. Se utilizó del paquete Microsoft Office las aplicaciones Microsoft Excel para el almacenamiento y procesamiento los datos primarios, cálculo de porcentaje e indicadores de calidad; el Microsoft Word para la redacción del informe final y el Microsoft Power Point para la exposición del informe final. El SPSS versión 15.0 para Windows fue utilizado para el cálculo del Alfa de Cronbach y el Test de correlación de Spearman. Los resultados se presentaron en tablas y gráficos.

Análisis y discusión.

Caracterización del instrumento de evaluación.

Al concluir el programa de la asignatura Metodología de la Investigación y Estadística en el curso 2010-2011 se aplicó el examen final escrito al que se presentaron 276 estudiantes en el municipio Holguín.

La información aportada por la guía utilizada para la descripción del examen, la base de datos confeccionada para el procesamiento estadístico y la entrevista aplicada a los expertos, permitieron la obtención de los resultados que se detallan a continuación:

El examen final estuvo compuesto por cinco preguntas (Anexo 2) con un total de 32 ítems de tipo ensayo, predominaron los ítems de respuestas cortas, 4 fueron de completamiento. El grupo de expertos consideró que es común en esta asignatura la elaboración de exámenes con preguntas de tipo ensayo en las que los estudiantes pueden expresar sus ideas con claridad.

La autora del presente trabajo considera importante tener en cuenta en la confección de los exámenes que las preguntas de ensayo limitan el número de áreas a explorar, la objetividad es difícil de conseguir, su calificación consume demasiado tiempo y sus resultados son poco confiables.

Autores como Moreno Montañez M, Carrazana Lee A, Abreu Reyes DR, analizaron en sus investigaciones instrumentos evaluativos con formato mixto, resultados que evidencian la tendencia de elaboración de exámenes con este formato en ciencias médicas.^{2, 12, 65.} Criterio que se debe seguir para la elaboración de futuros exámenes en la asignatura Metodología de la Investigación y Estadística.

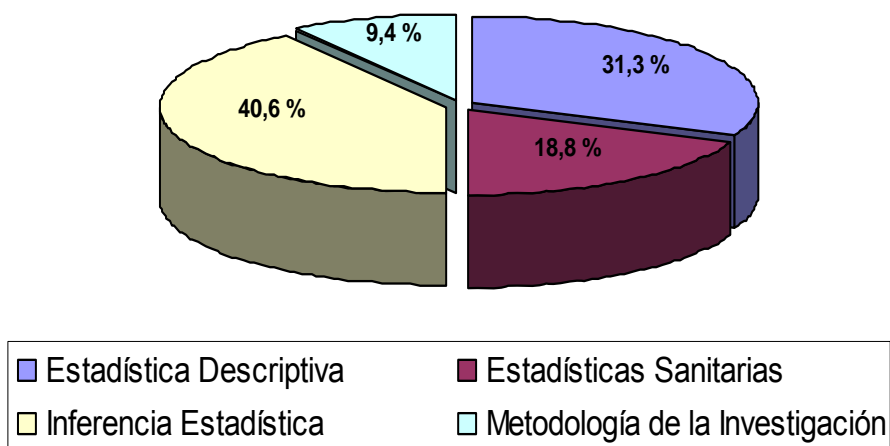
El examen en estudio exploró contenidos esenciales de todos los temas del programa de la asignatura que incluye Estadística descriptiva, Estadísticas Sanitarias, Introducción a la inferencia estadística y Metodología de la investigación. El cuadro 1 y gráfico1 representan el porcentaje de ítems aplicados en cada área temática.

Cuadro1. Ítems por área temática.

Temas	Ítems	%
Estadística Descriptiva	10	31,25
Estadísticas Sanitarias	6	18,75
Inferencia Estadística	13	40,63
Metodología de la Investigación	3	9,37
Total	32	100,00

Fuente: Examen Final de Metodología de la Investigación y Estadística.

Gráfico 1
Porcentaje de ítems según área temática



Fuente: Cuadro 1.

El tema estadística inferencial fue el más abordado con 40,6 % (13) de los ítems evaluados, seguido por el de estadística descriptiva con 31,3 % (10) y el menos explorado fue el de metodología de la investigación con 9,4 % (3).

El proceso de calificación del examen se realizó en la sede central por el colectivo de profesores del departamento de Informática Médica que impartían la asignatura, apoyados en la clave de calificación. Se consideraron desaprobados (suspensos) los que obtuvieron calificación de dos y aprobados los que obtuvieron calificación de tres, cuatro o cinco.

El cuadro 2 muestra la cantidad de estudiantes según la calificación obtenida en el examen, se aprecia que de los 276 estudiantes examinados, 55 (20 %)

suspendieron y solo 28 (10 %) alcanzaron la máxima calificación. La promoción obtenida en el municipio Holguín fue 80 %.

Cuadro 2. Resultados del examen aplicado.

Calificación	Cantidad estudiantes	%
5	28	10,0
4	90	33,0
3	103	37,0
2	55	20,0
Total	276	100

Fuente: Examen Final de Metodología de la Investigación y Estadística.

Es significativo destacar que la mayor cantidad (57 %) de los estudiantes presentados obtuvieron calificaciones de 3 o 2. Para Álvarez de Zayas³⁷ "la evaluación está determinada por los objetivos..., la evaluación influye en lo que enseñamos y determina de manera directa en lo que los educandos aprenden". El grupo de expertos coincidió con lo anterior al expresar que según los resultados obtenidos en este examen, algunos objetivos del programa no fueron vencidos en su totalidad por los estudiantes durante el semestre.

La pregunta uno estuvo formada por siete ítems, con predominio del tipo ensayo de respuesta corta. Exploró la clasificación clínico epidemiológica de un estudio, elaboración de problema científico y del objetivo general dada una situación de salud, clasificación de variables según escala de medición y tamaño de la muestra.

La pregunta dos con un ítem de desarrollo, dos de respuestas cortas y cuatro de completamiento, exploró los contenidos de cálculo de medidas de resumen para datos cualitativos e interpretación de una tabla estadística.

La pregunta tres exploró el contenido de estadísticas sanitarias con tres ítems de respuestas cortas y tres de desarrollo sobre el cálculo e interpretación de tasas de natalidad, mortalidad y morbilidad.

La pregunta cuatro comprobó la distribución Ji cuadrado y prueba de hipótesis de homogeneidad para el análisis e interpretación de una situación de salud donde

se aplique la prueba de hipótesis. Estuvo formada por dos ítems de desarrollo y tres de respuestas cortas.

La pregunta cinco abarcó contenido referente a correlación y regresión lineal con cuatro ítems de desarrollo y tres de respuestas cortas.

Correspondencia del fondo de tiempo y los ítems explorados en el examen.

La correspondencia del tiempo dedicado según el plan calendario (P_1) de la asignatura a cada tema y la cantidad de ítems que exploran estos contenidos en el examen, se muestran en el cuadro 3.

Cuadro 3. Fondo de tiempo del programa y presencia de ítems en el examen según temas.

Temas	Fondo de tiempo		Presencia en el examen	
	Horas	%	Ítems	%
Estadística descriptiva	16	26,67	10	31,25
Estadísticas sanitarias	12	19,99	6	18,75
Inferencia estadística	16	26,67	13	40,63
Metodología de la investigación	16	26,67	3	9,37
Total	60	100,00	32	100,00

Fuente: Programa de la asignatura y examen final.

En el cuadro anterior se observa que el programa de la asignatura tiene un fondo de tiempo de 60 horas distribuidas en cuatro temas, tres de ellos con 16 horas clases cada uno que representa 26,67 % respectivamente. Por otra parte de los 32 ítems explorados en el examen 13 que representan 40,63 % corresponden al tema Inferencia estadística y solo 3 que representan 9,37 % al tema Metodología de la investigación.

Se observó una marcada diferencia entre la cantidad de ítems en estos temas cuando ambos presentan el mismo tiempo en el plan calendario, evidencia de que no existió correspondencia entre el porcentaje de ítems explorados en los diferentes temas y el porcentaje del tiempo dedicado a cada uno en el plan calendario.

Para corroborar dichos resultados se calculó el Test de correlación de Spearman y se obtuvo $r = 0,25$ IC 95 %, reflejo de una débil relación entre ambas variables.

Al considerar que debe existir una estrecha relación entre los objetivos que se imparten y los que se evalúan, dado por los contenidos explorados en el examen, se muestra que en el instrumento de evaluación aplicado no se cumplió proporcionalmente.

Resultados que entran en contradicción con lo planteado por Olga de Landaluce³⁹ cuando expresa “la relación entre objetivos y contenidos se comprobará mediante la evaluación” y con los obtenidos por Enrique Santienteban²³ al comprobar que en su estudio existió correspondencia entre fondo de tiempo dedicado a cada temática y el número de incisos evaluados en el examen en cada una de ellas.

Por las características de la carrera, el programa de la asignatura está muy saturado, se unen en ella dos ciencias complejas (Metodología de la Investigación y Estadística) e importantes que contribuyen al desempeño del médico general en el desarrollo de las funciones declaradas en el perfil del egresado.

La autora comparte el criterio de los expertos que el tema metodología de la investigación es uno de los más difíciles y el sistema de evaluación del tema incluye un seminario evaluativo (el único en el programa), razones que pudieran haber incidido en su menor representatividad en el examen y asume que puede ser más explorado si se confeccionan preguntas de tipo test objetivo que permitan evaluar mayor cantidad de contenido.

Es importante que el colectivo de profesores tenga presente que el referido tema le aporta al futuro médico elementos necesarios para resolver problemas de salud presentes en su desempeño laboral a través de investigaciones biomédicas.

Calidad del instrumento aplicado

Para el análisis de la calidad del examen de Metodología de la Investigación y Estadística aplicado, se tuvo en cuenta la dificultad, discriminación y consistencia interna.

El índice de dificultad es la proporción de personas que responden correctamente un reactivo de una prueba. Entre mayor sea esta proporción, menor será su dificultad. Se trata de una relación inversa: a mayor dificultad del ítem, menor será su índice.¹⁴ En el anexo 3 se relaciona la pregunta, tema e índice de dificultad de cada ítem del examen.

La pregunta 1 evaluó objetivos referentes a los temas estadística descriptiva, inferencial y metodología de la investigación, el resto de las preguntas evaluaron contenidos de un solo tema, las preguntas 4 y 5 exploraron el tema de inferencia estadística.

De los 32 ítems explorados, se evidencia que 16 (50 %) clasifican de dificultad media cuyos índices se encuentran entre 0,51 y 0,80; seis clasifican como fáciles con índice mayor que 0,90, todos referentes al tema estadística descriptiva. Seis como medianamente fáciles con índices entre 0,81 y 0,90 referente a los temas estadística descriptiva, estadísticas sanitarias e introducción a la inferencia estadística. Dos clasifican como medianamente difíciles con valores de 0,40 a 0,50 y dos como difíciles con índices menores de 0,40, los cuatro referentes al tema inferencia estadística.

Resultados que difieren de los obtenidos por Abreu Reyes⁶⁵ al plantear que en su estudio no se encontraron ítems difíciles ni medianamente difíciles.

En el cuadro 4, se refleja la distribución de los ítems esperados y observados según su grado de dificultad de acuerdo a la escala establecida por el criterio de expertos, con un comportamiento diferente en el porcentaje de los ítems fáciles entre el valor esperado (5 %) y el real (20 %), y en los medianamente difíciles donde disminuyó considerablemente el real (14 %) respecto al esperado (20 %), los ítems fáciles y difíciles (niveles extremos de dificultad) alcanzan el 25 % de los explorados.

Los resultados obtenidos difieren de lo planteado por Backhoff¹⁵ como valor esperado teóricamente en un examen, cercano al 10 % y del obtenido en los reactivos del EXHCOBA de solo 12 %. El resto de los niveles de dificultad se comportaron de modo similar en cuanto al valor esperado y observado.

Cuadro 4. Ítems del examen según grado de dificultad esperado y observado.

Grado de dificultad	Ítems esperados		Ítems observados	
	Cantidad	%	Cantidad	%
Fáciles	2	5	6	20
Medianamente fáciles	6	20	6	20
De dificultad media	16	50	16	50
Medianamente difíciles	6	20	2	5
Difíciles	2	5	2	5

Fuente: Examen Final de Metodología de la Investigación y Estadística.

Este resultado evidencia que no existió gran correspondencia entre los valores esperados y observados en el instrumento objeto de estudio. La aplicación del coeficiente de correlación de Spearman ($r = 0,55$ IC 95 %) permitió corroborar lo expuesto.

La autora considera que no se realizó la planificación del examen con la profundidad requerida y esto pudo ser la causa de la diferencia de los valores porcentuales esperados y observados de los incisos fáciles y medianamente difíciles.

Al calcular el índice de dificultad por temáticas y examen representado en el cuadro 5, se observa que el instrumento aplicado en la asignatura Metodología de la Investigación y Estadística presentó una dificultad de 0,74, considerado de dificultad media. Con una ligera disminución de la dificultad a la obtenida por Backhoff en sus estudios, entre 0,5 y 0,6.¹⁵

Cuadro 5. Dificultad del instrumento evaluativo por área temática y examen.

Área temática	Índice de dificultad	Tipo de dificultad
Estadística Descriptiva	0,87	Medianamente fácil
Estadísticas Sanitarias	0,83	Medianamente fácil
Inferencia Estadística	0,61	Dificultad media
Metodología de la Investigación	0,63	Dificultad media
Examen	0,74	Dificultad media

Fuente: Procesamiento estadístico n=276

Los temas de mayores dificultades fueron el de inferencia estadística y metodología de la investigación con índice de 0.61 y 0.63 respectivamente, clasificados de dificultad media. Coincidieron en ser los temas donde existió mayor diferencia entre el porcentaje de ítems a explorar y el porcentaje de tiempo dedicado a la temática en el programa.

Al analizar el índice de dificultad por preguntas en el cuadro 6, se destaca el predominio de las preguntas de dificultad media, en correspondencia con la dificultad general del examen de 0,74. Resultaron las preguntas dos y tres las de menores dificultades, clasificadas como medianamente fáciles, con valores de 0,86 y 0,83 respectivamente.

La pregunta cuatro fue la más difícil, clasificada de dificultad media, con un índice de 0,58. No hubo en el examen preguntas fáciles, resultado que difiere de los obtenidos en los estudios realizados por Moreno Montañez M, Pérez Tapia JH., Argudín Somonte E y Abreu Reyes D.^{2, 14, 18, 60.} en los que al menos una pregunta clasificó como fácil.

Cuadro 6. Índice de dificultad por preguntas y examen.

Pregunta	Dificultad	Tipo de dificultad
Pregunta 1	0,76	Dificultad media
Pregunta 2	0,86	Medianamente fácil
Pregunta 3	0,83	Medianamente fácil
Pregunta 4	0,58	Dificultad media
Pregunta 5	0,61	Dificultad media
Examen	0,74	Dificultad media

Fuente: Procesamiento estadístico n=276

La pregunta dos fue la de mayor índice de dificultad y por tanto la más fácil del examen, referente al tema de estadística descriptiva, contenido que se evaluó en el trabajo de control y en la prueba intrasemestral. La autora considera que al existir una mayor sistematización de los contenidos por parte de los estudiantes, lograron un mayor desarrollo de habilidades sobre el tema, aunque fueron evaluadas esencialidades no exploradas en las evaluaciones anteriores.

Es importante destacar que cuatro de los siete ítems evaluados en esta pregunta fueron de completamiento, exploraron la habilidad de cálculo de técnicas de la estadística descriptiva, como algunas de las medidas de resumen para datos cualitativos, a nivel reproductivo y coincidieron en ser clasificados como fáciles según su índice de dificultad. Solo un ítem fue de desarrollo y resultó el de mayor dificultad en la pregunta.

La pregunta cuatro, por el contrario, resultó la más difícil del examen al mostrar menor índice de dificultad, compuesta por cinco ítems; dos de ellos de desarrollo, clasificados como difíciles según su grado de dificultad. El grupo de expertos y la autora consideraron que fueron ítems de aplicación e interpretación de la prueba de hipótesis no paramétrica de Ji cuadrado, que necesitan razonamiento del estudiante para poder identificar el tipo de prueba, aplicarla e inferir los resultados obtenidos, y que a pesar de ser exploradas de forma similar en evaluaciones sistemáticas y parciales durante el proceso, no se alcanzó el nivel de asimilación requerido por gran parte de los estudiantes.

La autora opina que este es un contenido de mucha aplicación en investigaciones biomédicas y pudiera ser que el colectivo de profesores y estudiantes le dedicaron mayor tiempo a la sistematización de otro contenido del tema que consideraron de mayor dificultad.

Discriminación

Según Backhoff “Si la prueba y un ítem miden la misma habilidad o competencia, es de esperar que quien tuvo una puntuación alta en todo el test deberá tener altas probabilidades de contestar correctamente el ítem. También debemos esperar lo contrario....”¹⁵

El anexo 4 relaciona los resultados obtenidos al calcular el coeficiente de correlación del punto biserial (r_{pbis}) en cada ítem del examen en estudio. De los 32 ítems explorados, 20 presentan un excelente poder discriminativo, con valores mayores que 0,35, ocho discriminan bien, con valores entre 0,26 y 0,35, tres regular, con valores entre 0,15 y 0,25, y uno discrimina pobremente, entre 0 y 0,14.

Al agrupar los ítems según su coeficiente de correlación del punto biserial en el cuadro 7, se destaca que discriminan de forma excelente 62,5 % de ellos y solo 3,1 % presentaron pobre discriminación.

Cuadro 7. Porcentaje de ítems según coeficiente de discriminación r_{pbis} .

Discriminación	Cantidad de ítems.	% de ítems.
Excelente	20	62,5
Buena	8	25,0
Regular	3	9,4
Pobre	1	3,1
Total	32	100

Fuente: Anexo 4.

Los tres ítems que discriminan regular (9,4 %) y el que tiene un poder discriminativo pobre (3,1 %) coincidieron en ser clasificados como fáciles según su índice de dificultad, corresponden al tema de estadística descriptiva, fueron elaborados de tipo completamiento, con poca profundidad y sus resultados no discriminan entre los estudiantes con calificaciones altas y bajas.

Según algunos autores para rangos similares de discriminación la conducta a seguir es revisarlos a profundidad o descartarlos definitivamente y consideran que un buen ítem debe discriminar entre aquellos que obtuvieron buenas calificaciones en la prueba y aquellos que obtuvieron bajas calificaciones.^{12, 15.}

Es criterio de la autora que los ítems fáciles que no discriminan adecuadamente no deben ser incluidos de esta forma en posteriores exámenes, deben ser revisados con cautela para lograr reelaborarlos con mayor complejidad.

Coincide con lo planteado por María Eugenia Ponce de León⁶⁶ "Si el mayor número de alumnos que contestaron correctamente el reactivo tienen un alto rendimiento en el examen, el reactivo discrimina positivamente; si por el contrario fue contestado por un mayor número de alumnos con bajo rendimiento en el examen, el reactivo discrimina negativamente y deberá ser eliminado".

Estos resultados evidencian dificultades en la planificación y elaboración de instrumentos evaluativos con diferentes formatos, que exploren con la profundidad requerida la mayor cantidad de objetivos posibles de la asignatura, es necesario trabajar la línea de la evaluación del aprendizaje en las actividades metodológicas del departamento de Informática Médica y profundizar en el rediseño de los instrumentos de evaluación, que favorecen la eficiencia del proceso docente educativo.

Autores como Carrazana¹² y Backhoff¹⁵ plantean que el coeficiente de correlación del punto biserial es una medida que combina la relación entre el criterio del reactivo y el nivel de dificultad. En el examen en estudio existe coincidencia con lo anterior, se evidencia que los ítems que discriminaron regular y pobre resultaron fáciles según su índice de dificultad.

Al agrupar el coeficiente de correlación del punto biserial por preguntas del examen, se confeccionó el cuadro 8 y el gráfico 2. Se observa que los valores de este indicador oscilaron entre 0,40 y 0,68, evidencian la discriminación excelente ($r_{pbis} > 0,35$) de todas las preguntas del examen en estudio, resultados que coinciden con lo planteado por Backhof¹⁵ como las cifras promedio que se buscan en un examen para valorar la capacidad de predicción de un instrumento, y con los obtenidos por autores como Carrazana¹¹ al considerar en su estudio que las preguntas del examen miden el mismo conocimiento o habilidad que el examen

total y todas las preguntas permitieron predecir que el estudiante que salió bien en el examen debe haber salido bien en cada una de las preguntas.

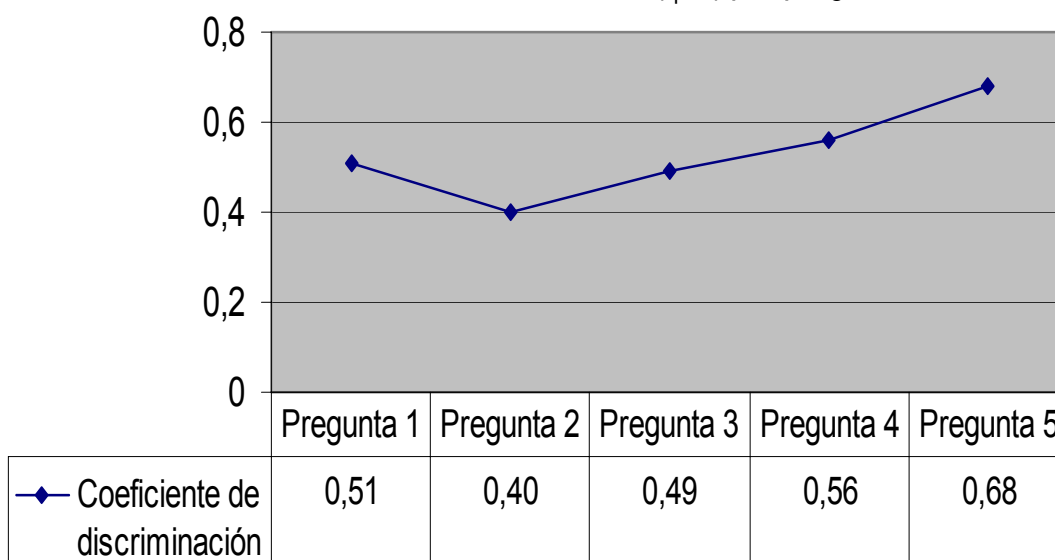
Cuadro 8. Coeficiente de discriminación correlación del punto biserial (r_{pbis}) por preguntas.

Pregunta	R_{pbis}	Discriminación
Pregunta 1	0,51	Excelente
Pregunta 2	0,40	Excelente
Pregunta 3	0,49	Excelente
Pregunta 4	0,56	Excelente
Pregunta 5	0,68	Excelente

Fuente: Procesamiento estadístico n=276

Gráfico 2

Coeficiente de discriminación (r_{pbis}) por preguntas.



Fuente: Cuadro 8.

Al realizar un análisis exhaustivo de cada pregunta se observó que el coeficiente de discriminación más alto corresponde a la pregunta cinco que evalúa los objetivos de correlación y regresión lineal, seguido por la pregunta cuatro que evalúa prueba de hipótesis de homogeneidad, pertenecientes ambas preguntas al tema de inferencia estadística.

La pregunta dos aunque discriminó de forma excelente presentó el menor coeficiente de discriminación (0,40) del examen, en ella se exploran objetivos de estadística descriptiva y coincide en ser la de mayor índice de dificultad, caracterizada como medianamente fácil. A pesar de tener dos ítems que discriminaron de forma excelente, el resto lo hace regular o pobre.

La autora coincide con los expertos que en esta pregunta existió predominio de ítems con nivel de asimilación reproductivo, que requieren de cálculos muy elementales para su solución y del modo en que fueron elaborados no le permitieron al médico en formación aplicar los métodos que proporciona la estadística descriptiva para resumir y presentar información biomédica.

Confiabilidad.

El coeficiente de confiabilidad Alfa de Cronbach es utilizado para valorar la coherencia interna global de la escala.^{6, 63.}

El cuadro 9 presenta los resultados del coeficiente Alfa de Cronbach para cada una de las preguntas y para el examen. Se observa que el examen tiene una coherencia interna de 0,65, demuestra que su confiabilidad es aceptable, resultado similar a los establecidos por la Comisión Nacional de Evaluación de la Competencia y el Desempeño del Ministerio de Salud Pública que propone valores superiores a 0.60.⁶²

Cuadro 9. Coeficiente Alfa de Cronbach por preguntas y examen.

Preguntas	Cronbach
Pregunta 1	0,70
Pregunta 2	0,59
Pregunta 3	0,58
Pregunta 4	0,61
Pregunta 5	0,53
Examen	0,65

Fuente: Procesamiento estadístico n=276

Al eliminar la pregunta uno el coeficiente alfa de Cronbach aumenta ligeramente respecto al obtenido en el examen, evidencia que aumenta la fiabilidad del

examen. Esta pregunta no aportó lo suficiente a la coherencia interna del instrumento, debe ser revisada para próximas evaluaciones.

La autora considera que a pesar de explorar contenidos de tres temas no se logró una adecuada interrelación en el momento de planear el examen.

Las restantes preguntas, al presentar valores inferiores al obtenido en el examen, se considera que contribuyen de igual manera a la confiabilidad del instrumento.

- Caracterización del instrumento de evaluación.

- **Descripción del examen final.**

- 1.- El examen objeto de estudio tuvo una duración de tres horas, fue estructurado con 32 ítems agrupados en cinco preguntas, con predominio de los ítems de ensayo de respuestas cortas.
- 2.- Se exploraron contenidos de los cuatro temas del programa analítico de la asignatura que dan salida al cumplimiento de objetivos generales del Médico General plasmados en el perfil del egresado de Ciencias Médicas, el tema menos representado fue Metodología de la Investigación.
- 3.- Aprobaron 221 estudiantes (80 %) de los 276 presentados al examen en el municipio Holguín, con predominio de calificación regular (3) y bien (4).

- **Relación entre fondo de tiempo y la cantidad de ítems.**

- 1.- No existió correspondencia entre el tiempo dedicado a cada tema según programa y la cantidad de ítems explorados en el examen de cada tema, resultado que se corroboró a través del Test de correlación de Spearman.

- **Calidad del instrumento de evaluación aplicado.**

Índice de dificultad.

- 1.- No existió correspondencia en la distribución del nivel de dificultad esperado y real, corroborado mediante el Test de correlación de Spearman.
- 2.- Las preguntas 4 y 5 resultaron las más difíciles, con un índice de dificultad de 0,58 y 0,61 respectivamente y la pregunta 2 la más fácil con valor de 0,86.

- 3.- El tema de Inferencia Estadística resultó el más difícil y el de Estadística Descriptiva el más fácil con una dificultad de 0,61 y 0,87 respectivamente.
- 4.- De forma general el examen presentó dos preguntas medianamente fáciles y tres de dificultad media. La dificultad del instrumento fue 0,74, considerado como un examen de dificultad media.

Coefficiente de correlación del punto biserial.

- 1.- Este indicador demostró que un ítem discriminó pobre y tres regular, coincidieron en ser clasificados fáciles según su dificultad. Valorados por el grupo de expertos y la autora de este trabajo con dificultades en su elaboración y nivel de asimilación reproductivo.
- 2.- El coeficiente de correlación del punto biserial por preguntas fue superior a 0,35, resultados que evidencian la discriminación excelente de cada una y del examen como un todo, lo que permitió predecir que el estudiante que salió bien en el examen debe haber salido bien en cada una de las preguntas.

Alfa de Cronbach.

- 1.- Al realizar el análisis de la consistencia interna del examen mostró un valor de 0.65, considerado con una confiabilidad aceptable y concuerda con el coeficiente que se propone en las Indicaciones Metodológicas para el Planeamiento, Diseño y Control de la Calidad de los Exámenes Escritos para la Evaluación de la Competencia y el Desempeño Profesional.
- 2.- Al eliminar la pregunta 1 se observa un ligero incremento de este indicador respecto al resultado del examen, considerada por el grupo de expertos y la autora con dificultad en su elaboración.

Conclusiones.

- Fueron evaluados contenidos de todos los temas de la asignatura, expresado con menor frecuencia en el examen el de Metodología de la Investigación. Predominaron los ítems de ensayo en su variante de respuestas cortas. Más de la mitad de los estudiantes obtuvieron bajas calificaciones.

- No existió correspondencia entre el tiempo dedicado a cada tema y la cantidad de ítems explorados en el examen.

- La valoración de los indicadores de calidad aplicados demostró que la cantidad de ítems fáciles superó el por ciento esperado. Los ítems que no discriminaron de forma adecuada coincidieron en ser clasificados fáciles, con dificultades en su elaboración y nivel de asimilación reproductivo. El examen presentó una dificultad media, excelente discriminación de todas las preguntas y una confiabilidad aceptable.

Recomendaciones.

- Contribuir a la correcta planeación de los instrumentos evaluativos a partir de los resultados obtenidos en esta investigación que favorezca la calidad del proceso docente educativo.
- Divulgar los resultados de la presente investigación en las diferentes formas de trabajo metodológico del departamento de Informática Médica.

Referencias bibliográficas.

1. Tünnermann, B. Conferencia regional sobre políticas y estrategias para la transformación de la educación superior en América Latina y el Caribe. La Habana, Cuba; 1996.
2. Moreno Montañez M. Caracterización del examen estatal escrito de la especialidad en Medicina General Integral. 1ra convocatoria. República Bolivariana de Venezuela. Mayo 2006. [Tesis para optar por el título de Máster en Educación Médica Superior]. República Bolivariana de Venezuela: ENSAP; 2007.
3. Caballero González JE. Análisis del perfil del egresado de la carrera de medicina en la Universidad Médica Cubana. Rev Educ Med Super [Internet]. 2002 ene-abril [citado 20 junio 2011]; 14(1): [Aprox. 5 p]. Disponible desde: www.bvs.sld.cu/revistas/ems/vol14_1_00/ems02100.htm
4. Morales I, Fernández B. La evaluación institucional en los centros de educación médica de Cuba. Rev Educ Med Sup [Internet]. 2005 [Citado 16 de feb 2011]; 19(3). Disponible desde: <http://scielo.sld.cu/scielo.php?script=sci-arttext&pid=S0864-21412005000300007&lng=es&nrm=iso&tlng=es>
5. Horruitiner Silva, P. La universidad cubana: el modelo de formación. La Habana: Editorial Félix Varela; 2008.
6. Salas Perea RS. La evaluación en la educación superior contemporánea. Biblioteca de Medicina. Volumen XXIV. UMSA, La Paz, Bolivia, 1998. En CD-ROM Maestría en Educación Médica, ENSAP; 2005.
7. Mesa Martín IM, Bravo Licourt MM, Linares Valdés H, Junco Arévalo J. Caracterización y resultados académicos de estudiantes de Medicina del Proyecto Policlínico Universitario. Educ Med Super [Internet]; 2011 jul.-set. [citado 22 noviembre 2011]; 25(3): [Aprox 6 p.]. Disponible desde: <http://scielo.sld.cu/scielo.php?script=sci-arttext&pid=S0864-21412011000300008&lng=es&nrm=iso&tlng=es>
8. Gimeno Sacristán, J. La pedagogía por objetivos: obsesión por su eficacia. Madrid: Morata; 1982.
9. Colectivo de Autores CECAM. Programa de estudio de Informática Médica. Instituto Superior de Ciencias Médicas; 2001.

10. Ministerio de Salud Pública. Comisión Nacional Carrera de Medicina. Perfeccionamiento del Plan de estudio de la Carrera de Medicina. Universidad de Ciencias Médicas de la Habana; 2010.
11. Mauri T, Coll C, Onrubia J. La evaluación de la calidad de los procesos de innovación docente universitaria. Una perspectiva constructivista. Rev Docencia Universitaria [Internet]. 2007 [citado 22 mayo 2011]; 1(1): [Aprox. 11 p]. Disponible desde: <http://revistas.um.es/redu/article/view/3341>
12. Carrazana Lee A, Salas Perea RS, Ruiz Salvador AK. Nivel de dificultad y poder de Discriminación del examen diagnóstico de la asignatura Morfofisiología Humana I. Educ Med Super [Internet]; 2011 ene-mar [citado 22 abril 2011]; 25(1): [Aprox. 12 p]. Disponible desde: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412011000100010&lng=es&nrm=iso&tlng=es
13. Pérez Espinosa L. Resultados de evaluación de competencia en la especialidad de MGI. Ciego de Ávila 2002. [Tesis para optar por el título de Máster en Educación Médica Superior]. La Habana: ENSAP; 2003.
14. Pérez Tapia JH , Acuña Aguila N, Arratia Cuela E.. Nivel de dificultad y poder de discriminación del tercer y quinto examen parcial de la cátedra de citohistología 2007 de la carrera de medicina de la UMSA. Rev Cuadernos [Internet]; 2008 [citado 22 abril 2011]; 53(2): [Aprox. 14 p]. Disponible desde: www.revistasbolivianas.org.bo/pdf/chc/v53n2/v53n2a03.pdf
15. Backhoff E., Larrazolo N., Rosas M. Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). Revista Electrónica de Investigación Educativa [Internet]; 2000. [citado 2 de octubre de 2011]; 2 (1): [Aprox. 14 p]. Disponible en: <http://redie.uabc.mx/vol2no1/contenido-backhoff.html>
16. Sánchez González MaG. Descripción del nivel de facilidad y poder de discriminación del examen de inferencia estadística en métodos estadísticos en Medicina veterinaria y Zootecnia. Facultad de Medicina Veterinaria y Zootecnia. México. Jornada de Educación Médica [Internet]; 2010. [Citado: 2 enero 2012]: [Aprox. 9 p]. Disponible desde: Disponible en: http://www.facmed.unam.mx/sem/jem2010/Abstracts/trabajos_orales.pdf
17. Caraballo Carballo RM, Rico Romero L, Lupiáñez Gómez JL. Análisis de los ítems de las evaluaciones autonómicas de diagnóstico en España: 2008-2009.

- Rev Iberoamericana de Educación Matemática [Internet]; 2011 Junio [citado 2 de octubre de 2011]; (26): [Aprox. 16 p]. Disponible desde: http://www.fisem.org/web/union/index.php?option=com_content&view=article&id=32&Itemid=3
18. Argudín Samonte E, Díaz Rojas P, Leyva Sánchez E. Índice de Dificultad del examen de Morfofisiología Humana I. Educ Med Super [Internet]; 2011 abr-Jun [citado 2 de octubre de 2011]; 25(2): [Aprox. 16 p]. Disponible desde: http://bvs.sld.cu/revistas/ems/vol25_2_11/ems07211.htm
 19. Santa Eulalia Mainegra JM. Caracterización de la Evaluación Frecuente durante el Encuentro en el Núcleo Docente, Modalidad Evaluación del Aprendizaje. Morfofisiología Humana III. Municipio Simón Bolívar, Estado Anzoátegui. Curso 2006- 2007. [Tesis para optar por el título de Máster en Educación Médica Superior]. La Habana: ENSAP; 2007.
 20. Marín Hernández D. Estructura y valoración de la calidad del examen final de Morfofisiología Humana I en el Estado Táchira. Curso 2005 – 2006. [Tesis para optar por el título de Máster en Educación Médica Superior]. La Habana: ENSAP; 2007.
 21. Zubizarreta MM. Planeamiento y calidad del examen estatal escrito de Licenciatura en Enfermería. Año 2001. [Trabajo para optar por el título de Máster en Educación Médica Superior]. La Habana: ENSAP; 2001.
 22. Díaz Roig I, de la Cruz González V, Cruz Ramírez C. El examen estatal escrito en la especialidad de Medicina General Integral. Rev Cubana Med Gen Integr [Internet]; 1998; [citado 2 de octubre de 2011]; 14(6): [Aprox. 6 p]. Disponible desde: http://bvs.sld.cu/revistas/mgi/vol14_6_98/mgi10698.htm
 23. Santiesteban Leyva E. Propuesta de metodología para el análisis de los resultados de los exámenes finales de las asignaturas de la disciplina Morfofisiología Humana del Programa Nacional de Formación de Medicina Integral Comunitaria en la República Bolivariana de Venezuela. [Tesis para optar por el título de Máster en Educación Médica Superior]. La Habana: ENSAP; 2007.
 24. Ilizástigui Dupuy F. Formación del médico general básico en Cuba. Material de estudio Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica. ENSAP: 2005.

25. Cires Reyes E, Vicedo Tomey A, Prieto Marrero G, García Hernández M. La investigación pedagógica en el contexto de la carrera de Medicina. *Educ Med Super Super* [Internet]; 2011 ene.-mar [citado 2 de febrero de 2012]; 25 (1): [Aprox. 7 p]. Disponible desde: bvs.sld.cu/revistas/ems/vol25_1_11/ems08111.htm
26. Álvarez Reyes S, Pérez Santos F, Otero Diéguez A, Tamayo Carralero I. La enseñanza de la estadística en las carreras de ingeniería: una propuesta de alternativas metodológicas para el tratamiento de la distribución de poisson. *Revista Pedagogía Universitaria* [Internet]. 2011 [citado 11 febrero 2012]; 16(2): [Aprox. 15 p]. Disponible desde: <http://cvi.mes.edu.cu/peduniv/base-de-datos/2011-vol.-xvi-no.-2/la-ensenanza-de-la-estadistica-en-las-carreras-de-ingenieria-una-propuesta-de-alternativas-metodologicas-para-el-tratamiento-de-la-distribucion-de-poisson>.
27. Hernández B, Rodríguez E. Universidad Médica y sociedad: su vinculación a la luz de la Informática Médica. *Rev Cub Inf Méd* [Internet]. 2005 [citado 11 noviembre 2011]; 1: [Aprox. 9 p]. Disponible desde: www.rcim.sld.cu/revista_7/articulo_7htm/univmesoc.htm
28. Greenes RA, Shortliffe EH. Medical informatics: An emerging academic discipline and institutional priority. *JAMA*; 1990.
29. Cruz Font J. Modelo didáctico para el desarrollo de competencias de la Informática Médica en la formación inicial del Médico General Básico. [Tesis en opción al grado científico de Doctor en Ciencia Pedagógicas]. Holguín; 2009.
30. Rojas Ochoa F. La enseñanza de la salud pública a estudiantes de Medicina en Cuba. *BVS* [Internet]. 2003 mar. [citado 5 febrero 2012]. [Aprox. 10p]. Disponible desde: http://www.bvs.sld.cu/revistas/ems/vol17_2_03/ems01203.htm
31. Pérez Pérez S. Importancia de la Informática Médica en la formación laboral del profesional de la salud. CD-R V Taller Nacional Científico Metodológico sobre Formación Laboral. Holguín, 2012.
32. Cruz Y. Propuesta para el perfeccionamiento del diseño curricular para la carrera de medicina en el programa de asignatura Informática Médica II [Tesis de Maestría]. Universidad de Holguín; 2008.

33. Colectivos de Autores CECAM. Programa de estudio de Informática Médica. Instituto Superior de Ciencias Médicas Habana; 1997
34. Gutiérrez Santiesteban E. Estrategia didáctica para la dinámica del proceso formativo de la Informática Médica. [Tesis en opción al grado científico de Doctor en Ciencias Pedagógicas]. Manzanillo; 2011.
35. Prieto Díaz V, Quiñones La Rosa I, Ramírez Durán G, et al. Impacto de las tecnologías de la información y las comunicaciones en la educación y nuevos paradigmas del enfoque educativo. Educ Med.Super [Internet]. 2011 ene-mar [citado 11 febrero 2012]; 25(1): [Aprox. 6 p]. Disponible desde: http://bvs.sld.cu/revistas/ems/vol25_1_11/ems09111.htm
36. Ojeda Cabrera A. Análisis crítico del programa de la asignatura Informática Médica II de Medicina. Rev cub inform méd [Internet]. 2010 [citado 25 marzo 2011]; (2): [Aprox. 16 p]. Disponible desde: http://www.rcim.sld.cu/revista_21/articulo_21.htm
37. Alvarez de Zayas C. "La dirección del proceso docente educativo". Tomado de: Hacia una escuela de Excelencia. La Habana: Editorial Academia; 1996.
38. Milán Licea MR, Fuentes González H, De la Peña Silva R. La evaluación como un proceso participativo. En: Ginoris Quesada O. Fundamentos didácticos de la Educación Superior cubana. La Habana: Editorial Félix Varela; 2009.
39. Landaluce Gutiérrez O. Pedagogía. Editorial Ciencias Médicas: La Habana; 2011.
40. Reglamento de Trabajo Docente y Metodológico. Ministerio de Educación Superior. Resolución 210; 2007.
41. Colectivo de autores. Pedagogía. La Habana: Editorial Pueblo y Educación; 1984.
42. Díaz Hernández RJ, Díaz Rodríguez L, Seiglie Rebolgar MA, García Enríquez, ME. Evaluación educativa en la formación laboral integral de estudiantes de Tecnología de la Salud, perfil Laboratorio Clínico. Edumecentro [Internet]. 2011 may.-ago. [citado 25 marzo 2011]; 3(2): [Aprox. 8 p]. Disponible desde: <http://www.edumecentro.sld.cu/pag/sumario3-2-2011.html>
43. Hernández Socarrás X. Evaluación del desempeño profesional de los estudiantes de la Licenciatura en Educación Preescolar. [Tesis en opción al grado científico de Doctor en Ciencias Pedagógicas]. Ciudad Habana; 2003.

44. Gómez León K. Propuesta de guía para la evaluación del docente en la dirección de la autopreparación. Municipio Iribarren, Lara, Venezuela. Venezuela. [Tesis para optar por el título de Máster en Educación Médica Superior]. La Habana: ENSAP; 2007.
45. Casanova M. El proceso de la evaluación educativa. Ciberdocencia [Internet]. 2003 Ago. [citado 7 febrero 2011]; [Aprox. 9 p]. Disponible en: <http://www.redesc.ilce.edu.mx/redescolar/cursosytalleres/evaluacion.htm>.
46. De la Garza E. La Evaluación Educativa. Rev Mex de Investigación Educativa [Internet]; 2004 oct-dic [citado 15 febrero 2011]; 9(23): [Aprox. 11 p]. Disponible desde: <http://www.redalyc.org/src/inicio/ArtPdfRed.jsp?iCve=14002302>
47. González Péra M. La evaluación del aprendizaje: tendencia y reflexión crítica. Rev Edu Med Sup [Internet]; 2001 [citado 15 diciembre 2011]; 15(1): [Aprox. 10 p]. Disponible desde:
48. Cerda Gutiérrez H. La evaluación como experiencia total. Logros-objetivos-procesos. Competencia y desempeño. Santa Fe de Bogotá: Editorial Magisterio; 2000.
49. IV Congreso CiberSociedad. Concepto, metodología y retos de la evaluación on line: Buenas prácticas para su fiabilidad. [Internet]. 2006 [Citado 4 diciembre 2011]; [Aprox. 9 p]. Disponible desde: <http://www.cibersociedad.net/congres>.
50. García Ramos L. Auto perfeccionamiento docente y creatividad. La Habana: Editorial Pueblo y Educación. 1996.
51. Achig L. La autoevaluación de las Instituciones de la Educación Superior: Guía para la autoevaluación de la Universidad de Cuenca. Ecuador. Rev Electrónica Iberoamericana de Calidad, Eficacia y Cambio en Educación. [Internet]. 2003 [Citado 20 septiembre 2011]; 1(2): [Aprox. 12 p]. Disponible desde: <http://redalyc.uaemex.mx/redalyc/pdf>
52. González Pérez M. La evaluación del aprendizaje. En: Ginoris Quesada O. Fundamentos didácticos de la Educación Superior cubana. La Habana: Editorial Félix Varela; 2009.
53. Segura Castillo, M. La perspectiva ética de la evaluación de los aprendizajes desde un enfoque constructivista. Rev Actualidades Investigativas en

- Educación [Internet]. 2007 ene-abr [Citado 27 septiembre 2011]; 7(1). [Aprox. 22 p]. Disponible desde: dialnet.unirioja.es/servlet/articulo?codigo=2273870
54. Ministerio de Salud Pública. Resolución Ministerial 15; 1988.
55. Burguete Ramos MD. Evaluación: concepto y métodos. Educare21 [Internet]. 2003 [Citado 27 septiembre 2011]; (1). [Aprox. 9 p]. Disponible desde: www.enfermeria21.com/pfw_files/cma/revistas/.../2003/.../1024.pdf
56. Salim R, Lotti de Santos M, Macchioni de Zamora N. Evaluación del examen: una herramienta para la innovación didáctica en el aula universitaria. Revista Iberoamericana de educación [Internet]. 2012 may. [Citado 2 junio 2012]; 59(1). [Aprox. 7p]. Disponible desde: www.rieoei.org/rie_contenedor.php?numero=4147&titulo=Evaluaci%F3n%20del%20examen:%20una%20herramienta%20para%20la%20innovaci%F3n%20did%E1ctica%20en%20el%20aula%20universitaria
57. Zabalza Miguel A. "Competencias docentes del profesorado universitaria. Calidad y desarrollo profesional". Material de estudio de la Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica: ENSAP; 2005.
58. Salas Perea RS. La calidad en el desarrollo profesional: Avances y desafíos. Material de estudio de la Maestría en Educación Médica. En CD-ROM Maestría en Educación Médica: ENSAP; 2005.
59. Molina Soldán EM^a. Instrumentos de evaluación en el proceso de enseñanza aprendizaje. Revista digital Investigación y Educación [Internet]. 2006 Agosto [Citado 25 febrero 2012]; 26(3): [Aprox. 7 p]. Disponible desde: http://www.google.com/cu/#sclient=psy-ab&hl=es-419&q=Molina+E.+Instrumentos+de+evaluaci%C3%B3n+en+el+proceso+de+ense%C3%B1anza+aprendizaje&og=Molina+E.+Instrumentos+de+evaluaci%C3%B3n+en+el+proceso+de+ense%C3%B1anza+aprendizaje&aq=f&aqi=&aql=&gs_l=hp.3...86183.93652.3.99681.7.7.0.0.4.811.5388.6-7.7.0...0.0.Wg1rRA4UIOQ&bav=on.2,or.r_gc.r_pw.,cf.osb&fp=291a3ae0bac43d46&biw=977&bih=558
60. Alfonso de Belmonte MR. Una nueva perspectiva integradora para evaluar aprendizajes significativos. Prog Planeamiento Educativo [Internet]. 2006. [consultado 6 Diciembre 2011], Disponible desde:

http://www.psi.uba.ar/academica/cursos_actualizacion/recursos/minmendoza.pdf

61. Díaz P. Introducción a la Investigación en Ciencias de la Salud. La Habana: Escuela Nacional de Salud Pública.; 2010.
62. Cuba. Ministerio de Salud Pública. Proyecto de Indicaciones Metodológicas para el Planeamiento, Diseño y Control de la Calidad de los Exámenes Escritos para Evaluación de la Competencia y el Desempeño Profesionales. La Habana: Escuela Nacional de Salud Pública; 2001.
63. García-Bellido R, González Such J, Jornet Meliá J.M. SPSS: Análisis de fiabilidad. innovaMIDE [Internet]. 2010. [consultado 11 noviembre 2011], Disponible desde: www.uv.es/innovamide/spss/fiabilidad.wiki
64. Córlica JL, Hernández Aguilar ML, Holloway Creed J, Dimos C. Propuesta de indicadores de calidad para la evaluación de la calidad de las preguntas en los exámenes de opción múltiple. I Congreso virtual de educación a distancia [Internet]. 2004. [consultado 6 febrero 2011]. Disponible desde: http://www.ateneonline.net/datos/91_03%20PONENCIA%20CORICA-ERNANDEZ-HOLLOWAY-DIMOU.pdf
65. Abreu Reyes D. Caracterización del examen teórico final de Morfofisiología Humana del perfil Laboratorio Clínico. Rev Educ Med Super [Internet]. 2011 ene-mar [consultado 6 febrero 2012]; 25(1): [Aprox. 9 p]. Disponible desde: http://www.scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412011000100003&lng=es&nrm=iso&tlng=es.
66. Ponce de León ME. Estudio comparativo de la aplicación de un examen diagnóstico en postgrado, utilizando dos formatos: por computadora e impreso. Gac Méd Méx. [Internet]. 2003 [citado 20 Ene 2011]; 139(4): [Aprox. 7 p]: Disponible desde: <http://www.medigraphic.com/espanol/e-htms/-e-gaceta/e-gm2003/e-gm03-4/em-gm034b.htm>

Anexos.

Anexo 1: Entrevista a expertos.

Estimado profesor:

Se realiza una investigación en el campo de la evaluación, necesitamos de su colaboración para obtener información acerca del tema, sus criterios serán importantes para el perfeccionamiento del proceso docente en la asignatura Metodología de la Investigación y Estadística,

Agradecemos su colaboración.

- 1- ¿Qué tipo de formato se utilizó en la elaboración de cada inciso del examen?
- 2- ¿Qué contenidos temáticos de la asignatura fueron explorados en cada uno?
- 3- ¿Cómo deben tomarse en consideración los resultados porcentuales de promoción para clasificar la dificultad de los incisos y preguntas del examen en: difíciles, medianamente difíciles, dificultad media, medianamente fáciles y fáciles?
- 4- ¿Qué pudo influir en el comportamiento de la dificultad alcanzada en cada tema?
- 5- ¿Qué pudo influir en que unos temas fueran más explorados que otros?

Anexo 2. Prueba Final de Metodología de la Investigación. Curso 2010 – 2011.

**Universidad de Ciencias Médicas de Holguín.
Prueba Final de Metodología y Estadística 2do año Medicina
Curso 2010-2011**

Nombre: _____ **Grupo:** _____ **No.** _____ **Nota** _____

En un estudio realizado sobre Enfermedades Cerebro Vasculares (ECV) en personas con edades comprendidas entre 50 y 70 años en un área de salud del municipio Holguín en el período febrero a julio del 2009, se seleccionaron al azar 40 pacientes con ECV y 60 pacientes que no presentaban ECV. Se controlan las siguientes variables: peso (Kg), glicemia (mmol/L), cardiopatía isquémica (si/no), la dieta alimentaría (buena, regular, mala); la obesidad (si/no), diabetes (si/no).

1. Según la situación problemática planteada anteriormente:

- a) Formule el problema de investigación.
- b) Elabore el objetivo general.
- c) Clasifique las variables obesidad, dieta alimentaria y glicemia; según la escala de medición.
- d) Clasifique el tipo de estudio que se desarrolló.
- e) Diga cómo es la muestra de este estudio según su tamaño.

2. En el siguiente cuadro se muestra la “Relación entre la Ocurrencia de Enfermedades Cerebro Vasculares, según el sexo. Área de Salud, 2009”.

ECV	Sexo			
	M		F	
	No	%	No	%
Si	28		12	21,8
No	17	37,8		78,2
Total		100	55	

- a) Complete el cuadro.
- b) Interprete los resultados presentados.
- c) Identifique e interprete la medida de resumen calculada según los datos de la tabla:

$$\text{Medida?} = \frac{28}{17} \approx 1,65$$

3. Del área de salud referida se conocen los siguientes datos:

- Población media total: 86 400
- Nacidos Vivos: 1580
- Cantidad diabéticos: 620
- Total de fallecidos: 112
- Defunciones menores de un año: 5
- Fallecidos por diabetes: 6
- Cantidad Hipertensos: 1300

Calcule e interprete:

- a) Tasa de mortalidad infantil.
- b) Tasa de letalidad de diabetes.
- c) Tasa de natalidad.

4. En el estudio se detectó que 29 de las 40 personas con ECV estudiadas presentaban diabetes. La literatura científica médica establece que el 75 % del total con ECV padecen de diabetes. Utilizando un nivel de significación del 1 % y la información anterior se obtuvo: el estadígrafo calculado es - 4,62 y los valores tabulados son $\pm 2,58$. Pruebe estos resultados utilizando la prueba de hipótesis correspondiente a los datos.

5. Algunos datos obtenidos de los pacientes con ECV según el peso y la glicemia son los siguientes: coeficiente de correlación muestral (r) es igual a 0,86 (Coeficiente de Pearson); la ecuación de regresión es $Y = - 31 + 0,68X$; así como el estadígrafo calculado es $T = 2,1$. $RA : \pm 1,979$; $\alpha = 0,05$.

- a) Cómo es la relación entre el peso y la glicemia a nivel muestral.
- b) Interprete el modelo de regresión dado.
- c) Estime e interprete el valor de la glicemia de un paciente cuyo peso es de 60 Kg.
- d) Pruebe cómo es la relación entre el peso y la glicemia a nivel poblacional.

Anexo 3. Pregunta, tema e índice de dificultad por ítem.

Preg. e ítem	Temáticas	Cantidad de suspensos	Dificultad del ítem	Tipo de dificultad
1.1	Metodología Investigación	131	0,53	Media
1.2		95	0,66	Media
1.3.1	Estadística Descriptiva	25	0,91	Fácil
1.3.2		25	0,91	Fácil
1.3.3		34	0,88	Medianamente fácil
1.4	Metodología Investigación	80	0,71	Media
1.5	Inferencia Estadística	71	0,74	Media
2.1.1	Estadística Descriptiva	7	0,97	Fácil
2.1.2		12	0,96	Fácil
2.1.3		6	0,98	Fácil
2.1.4		6	0,98	Fácil
2.2		56	0,80	Media
2.3.1		63	0,77	Media
2.3.2		124	0,55	Media
3.1.1	Estadísticas Sanitarias	39	0,86	Medianamente fácil
3.1.2		47	0,83	Medianamente fácil
3.2.1		30	0,89	Medianamente fácil
3.2.2		66	0,76	Media
3.3.1		29	0,89	Medianamente fácil
3.3.2		66	0,76	Media
4.1	Inferencia Estadística	186	0,33	Difícil
4.1.2		42	0,85	Medianamente fácil
4.1.3		71	0,74	Media
4.1.4		91	0,67	Media
4.1.5		190	0,31	Difícil
5.1	Inferencia Estadística	75	0,73	Media
5.2		159	0,42	Medianamente difícil
5.3.1		122	0,56	Media
5.3.2		140	0,49	Medianamente difícil
5.4.1		80	0,71	Media
5.4.2		77	0,72	Media
5.4.3		94	0,66	Media

Fuente: Examen Final de Metodología de la Investigación y Estadística.

Anexo 4. Coeficiente de discriminación punto biserial por ítems.

Ítem	Coeficiente de discriminación (r_{pbis})	Discriminación
Ítem 1	0,69	Excelente
Ítem 2	0,70	Excelente
Ítem 3	0,34	Buena
Ítem 4	0,29	Buena
Ítem 5	0,27	Buena
Ítem 6	0,50	Excelente
Ítem 7	0,56	Excelente
Ítem 8	0,20	Regular
Ítem 9	0,20	Regular
Ítem 10	0,17	Regular
Ítem 11	0,14	Pobre
Ítem 12	0,67	Excelente
Ítem 13	0,58	Excelente
Ítem 14	0,28	Buena
Ítem 15	0,52	Excelente
Ítem 16	0,66	Excelente
Ítem 17	0,43	Excelente
Ítem 18	0,63	Excelente
Ítem 19	0,48	Excelente
Ítem 20	0,67	Excelente
Ítem 21	0,92	Excelente
Ítem 22	0,29	Buena
Ítem 23	0,31	Buena
Ítem 24	0,42	Excelente
Ítem 25	0,90	Excelente
Ítem 26	0,54	Excelente
Ítem 27	0,73	Excelente
Ítem 28	0,30	Buena
Ítem 29	0,35	Buena
Ítem 30	0,59	Excelente
Ítem 31	0,62	Excelente
Ítem 32	0,66	Excelente

Fuente: Procesamiento estadístico

n=276