



COLEGIO DE POSTGRADUADOS

INSTITUCION DE ENSEÑANZA E INVESTIGACION EN CIENCIAS AGRÍCOLAS

CAMPUS MONTECILLO

POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA

ESTADÍSTICA

PRUEBA DE BONDAD DE AJUSTE PARA LA DISTRIBUCIÓN POISSON BASADA EN EL ÍNDICE DE DISPERSIÓN DE FISHER

FAUSTINO BRAVO HERNÁNDEZ

**TESIS
PRESENTADA COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:**

MAESTRO EN CIENCIAS

MONTECILLO, TEXCOCO, EDO. DE MÉXICO

2012

LA PRESENTE TESIS TITULADA **PRUEBA DE BONDAD DE AJUSTE PARA LA DISTRIBUCIÓN POISSON BASADA EN EL ÍNDICE DE DISPERSIÓN DE FISHER** REALIZADA POR EL ALUMNO **FAUSTINO BRAVO HERNÁNDEZ** BAJO LA DIRECCIÓN DEL CONSEJO PARTICULAR INDICADO, HA SIDO APROBADA POR EL MISMO Y ACEPTADA COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE

MAESTRO EN CIENCIAS
SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA
ESTADÍSTICA

CONSEJO PARTICULAR

CONSEJERO

DR. HUMBERTO VAQUERA HUERTA

ASESOR

DR. JOSE A. VILLASEÑOR ALVA

ASESOR

DR. BARRY C. ARNOLD

MONTECILLO, TEXCOCO EDO. DE MÉXICO, JULIO DE 2012.

RESUMEN

Considerando el Índice de Dispersión de Fisher se estableció una Prueba de Bondad de Ajuste, que permite afirmar con cierta confiabilidad que los datos en cuestión se distribuyen de acuerdo a la Distribución Poisson. Para la realización de ésta, se determinaron sus propiedades, mediante el uso de la Teoría Asintótica y la Simulación de Monte Carlo. Se concluye que la Potencia de la Prueba propuesta, tanto para los diferentes tamaños de muestra, como para los diferentes niveles de significancia, son mayores que las de las pruebas de bondad de ajuste consideradas como comparación en este estudio. Por lo tanto se demuestra que es mejor la Prueba de Bondad de ajuste para la distribución Poisson basada en el Índice de Dispersión de Fisher.

Palabras Clave:

Índice de Dispersión de Fisher, Prueba de Bondad de Ajuste, Distribución Poisson, Teoría Asintótica, Simulación de Monte Carlo, Potencia de la Prueba.

ABSTRACT

A goodness of fit test based on Fisher's Index of dispersion is developed to determine whether a given data set can be viewed as having a Poisson distribution. To determine the operating characteristics of this test, appropriate asymptotic distribution theory is discussed together with Monte Carlo simulation based results. The power of the proposed test for various sample sizes and various levels of significance is shown to compare favorably with the power of several other available goodness of fit tests. As a consequence, use of the test based on Fisher's index of dispersion can be recommended to determine whether a given data set is well fitted by a Poisson model.

Key words:

Fisher's dispersion index, goodness of fit test, Poisson distribution, asymptotic theory, Monte Carlo simulation, power of a test.

AGRADECIMIENTOS

Al **Consejo Nacional de la Ciencia y Tecnología (CONACyT)**, por otorgarme el apoyo económico durante la realización de mis estudios de Maestría.

Al **Colegio de Postgraduados**, en especial al **Postgrado de Socioeconomía, Estadística e Informática - Estadística**, por haberme brindado las condiciones necesarias para concluir una etapa más de mi formación profesional.

Al **Dr. Humberto Vaquera Huerta**, por dirigir acertadamente la presente investigación y orientarme correctamente en la toma de decisiones.

Al **Dr. José A. Villaseñor Alva**, por sus sugerencias a la realización de la presente investigación

Al **Dr. Barry C. Arnold**, por su conocimiento y experiencia plasmados en el tema y determinantes en la culminación de esta investigación.

DEDICATORIA

A **Dios**, por permitirme lograr mis metas con éxito, amor y salud. Y no dejarme rendir ante las adversidades de la vida.

A mis padres Sra. **Teresa Hernández Estrada**⁺ y Sr. **Apolonio Bravo Nava**, quienes me dieron la vida y me educaron con principios y valores determinantes para la toma de decisiones en las diferentes etapas de mi vida.

A mi amada esposa **Catalina Delgado Aguilera**, quien es mi gran amor y que gracias a su madurez me ha abierto los ojos para darme cuenta lo valiosa que es como pareja, pues desde que esta a mi lado siempre cuento con su apoyo incondicional tanto en los momentos buenos y malos de mi vida.

A mis queridas hijas **Marlene y Jessica**, quien son mi inspiración y la razón principal para romper los obstáculos que se me presentan y con ello alcanzar mis metas de mi vida.

A mis hermanos **Carmen, Bonfilio, Pedro, Delfina, Víctor y Rosa**, por sus acertados consejos. Y a sus respectivas familias por la vida compartida que me han proporcionado.

CONTENIDO GENERAL

	Pág.
RESUMEN_____	i
ABSTRACT_____	ii
AGRADECIMIENTOS_____	iii
DEDICATORIA _____	iv
CONTENIDO GENERAL_____	v
CONTENIDO DE TABLAS_____	viii
CONTENIDO DE GRAFICAS_____	x
1.- INTRODUCCIÓN_____	1
1.1.- ANTECEDENTES_____	2
1.2.- PLANTEAMIENTO DEL PROBLEMA_____	5
1.3.- OBJETIVOS_____	6
2.- MARCO TEORICO_____	7
2.1.- DISTRIBUCIONES DE VARIABLES DE CONTEO_____	7
2.1.1.- DISTRIBUCIÓN POISSON_____	7
2.1.2.- DISTRIBUCIÓN BINOMIAL_____	9
2.1.3.- DISTRIBUCIÓN BINOMIAL NEGATIVA_____	10

2.2.- PRUEBA DE BONDAD DE AJUSTE_____	11
2.2.1.- PROBLEMA_____	11
2.2.2.- DEFINICIÓN_____	12
2.3.- ALGUNAS DEFINICIONES RELACIONADAS CON LAS PRUEBAS DE HIPÓTESIS_____	13
2.4.- EL METODO DELTA PARA CALCULAR DISTRIBUCIONES ASINTÓTICAS_____	15
2.4.1.- CASO UNIVARIADO_____	15
2.4.2.- CASO BIVARIADO _____	16
3.- PRUEBA DE BONDAD DE AJUSTE DEL INDICE DE DISPERSIÓN I_____	17
3.1.- DISTRIBUCIÓN DE LA ESTADÍSTICA DE PRUEBA_____	17
3.1.1.- OBTENCIÓN EN FORMA ANALÍTICA_____	18
3.1.2. - OBTENCIÓN EN FORMA EMPIRICA_____	21
3.2.- CONSTRUCCIÓN DE LA PRUEBA DE BONDAD DE AJUSTE_____	25
3.3.- TAMAÑO DE LA PRUEBA_____	27
3.4.- POTENCIA DE LA PRUEBA_____	30
4.- COMPARACIÓN DE LA POTENCIA DE LA PRUEBA PROPUESTA VS PRUEBA DE BONDAD DE AJUSTE DEL ÍNDICE DE DISPERSIÓN DE FISHER_____	33

4.1.- PRUEBA DE BONDAD DE AJUSTE DEL ÍNDICE DE DISPERSIÓN DE FISHER	34
5.- EJEMPLO DE APLICACIÓN DE LA PRUEBA DE BONDAD DE AJUSTE PROPUESTA	37
6.- CONCLUSIONES	41
7.- RECOMENDACIONES	43
8.- BIBLIOGRAFIA	44
9.- ANEXOS	45
9.1.-ANEXO A	45
9.2.-ANEXO B	58

CONTENIDO DE TABLAS

Tabla 1.- Comparación de los cuantiles de la distribución de la estadística de prueba I^* se distribuye Poisson($\lambda = 1$), estimados con 20,000 muestras, por el método Monte Carlo vs los cuantiles de la distribución $N\left(1, \frac{2}{n}\right)$ _____	23
Tabla 2.- Ejemplo del tamaño de la prueba para la distribución Poisson($\lambda = 0.2$)_____	29
Tabla 3.- Ejemplo de la potencia de la prueba cuando la H_a es una Binomial con media 1 considerando los valores de la Normal Estándar _____	32
Tabla 4.- Ejemplo de la potencia de la prueba cuando la H_a es una Binomial con media 1 considerando los valores de la distribución Ji-cuadrada _____	36
Tabla 5.- Comparación de la potencia de la prueba propuesta vs prueba de bondad de ajuste del índice de dispersión de Fisher cuando H_a : Binomial con media 1_	36
Tabla 6.- Números de muertes debidas a patadas de caballos en el ejército Prusiano__	38
Tabla 7.- Conteos de excedencias del ozono arriba del umbral nivel 16_____	39
Tabla 8.- Cuantiles de la distribución $N\left(1, \frac{2}{n}\right)$ _____	47

Tabla 9.- Cuantiles de la distribución de la estadística de prueba I bajo H_0 de que x se distribuye Poisson($\lambda = 1$), estimados con 20,000 muestras, por el método Monte Carlo_____ 50

Tabla 10.- Cuantiles de la distribución de la estadística de prueba I bajo H_0 de que x se distribuye Poisson($\lambda = 2$), estimados con 20,000 muestras, por el método Monte Carlo_____ 53

Tabla 11.- Cuantiles de la distribución de la estadística de prueba I bajo H_0 de que x se distribuye Poisson($\lambda = 3$), estimados con 20,000 muestras, por el método Monte Carlo_____ 56

CONTENIDO DE GRAFICAS

Gráfica 1.- Comparación de los histogramas de la estadística de prueba I^* bajo H_0 de que x se distribuye Poisson($\lambda = 1$), estimados con 20,000 muestras, por el método Monte Carlo vs los histogramas de la distribución $N\left(1, \frac{2}{n}\right)$ _____	24
Gráfica 2.- Ejemplo probabilidad estimada del error tipo i de la prueba cuando la muestra proviene de la distribución Poisson($\lambda = 0.2$)_____	29
Gráfica 3.- Ejemplo de la potencia de la prueba cuando la H_a es una Binomial con media 1 considerando los valores de la Normal Estándar_____	32
Gráfica 4.- Histogramas de la distribución $N\left(1, \frac{2}{n}\right)$ _____	46
Gráfica 5.- Histogramas de la estadística de prueba I bajo H_0 de x se distribuye Poisson($\lambda = 1$), estimados con 20,000 muestras, por el método Monte Carlo_____	49
Gráfica 6.- Histogramas de la estadística de prueba I bajo H_0 de que x se distribuye Poisson($\lambda = 2$), estimados con 20,000 muestras, por el método Monte Carlo_____	52
Gráfica 7.- Histogramas de la estadística de prueba I bajo H_0 de que x se distribuye Poisson($\lambda = 3$), estimados con 20,000 muestras, por el método Monte Carlo_____	55

1.- INTRODUCCIÓN

En diferentes áreas de la ciencia, se tienen varios estudios científicos en los cuales surgen con frecuencia datos que se refieren a conteos, por ejemplo:

- En la agricultura: el número de pétalos en una flor, el número de flores en una planta, el número de plantas en cierta área dada, etc.
- En la medicina: el número de virus en una solución, el número de dientes deficientes en un individuo, el número de muertes por cáncer en un hogar, el número de casos de una enfermedad, etc.
- En la industria: el número de artículos defectuosos en un proceso industrial, etc.
- En la entomología: el número de insectos de una población, etc.
- Otros: el número de errores tipográficos, el número de accidentes en un periodo de tiempo, etc.

Como en este tipo de datos la distribución normal no es modelo adecuado, un problema que enfrenta el investigador una vez obtenido los datos de dichos estudios científicos, es saber con cierta seguridad de que modelo probabilístico provienen dichos datos, teniendo como antecedentes, por sus conocimientos propios del área o por referencias publicadas al respecto, que los datos podrían provenir de alguna de las distribuciones importantes, a saber, Poisson, Binomial, Binomial Negativa.

Ante este problema presentado surge la necesidad de plantear la siguiente interrogante: ¿Es posible obtener un criterio estadístico preciso, para seleccionar uno de los tres modelos mencionados, usando como base el índice de Fisher?

Con el propósito de contestar la pregunta mencionada, se han propuesto algunas estadísticas de prueba que son ampliamente usadas para probar si las observaciones de una muestra provienen de una distribución hipotetizada. En la siguiente sección se presenta un breve recuento de estas. Sin embargo cabe aclarar que esta investigación sólo aborda el caso de saber si en efecto los datos en cuestión provienen de una distribución Poisson.

1.1.- ANTECEDENTES

La estadística ji-cuadrada de Pearson (1900) es ampliamente usada para probar, si las observaciones de una muestra provienen de una distribución hipotetizada tales como Poisson, Binomial, Binomial Negativa, entre otras. Sin embargo, existen algunas dificultades en aplicar esta prueba, a saber:

- Se enfrenta con la decisión de que si es necesario agrupar los datos y en caso de ser así, definir cuantos grupos.
- La distribución asintótica de la estadística de Pearson podría no ser ji-cuadrada dependiendo de los estimadores empleados.
- La probabilidad requerida en el cálculo de la estadística de prueba puede conducir a cálculos más complicados, dependiendo de la forma de la distribución asumida.

Otra estadística para probar si los datos provienen de una distribución Poisson, contra la alternativa de que se originan de una distribución no Poisson, es el Índice de Dispersión Poisson, debido a Fisher (1950), definido por:

$$R = \frac{nS^2}{\bar{X}}, \text{ donde } S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \text{ y } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

la hipótesis H_0 es rechazada si $\left| \frac{R}{n} - 1 \right|$ es grande. Además bajo H_0 , la distribución asintótica de R es una distribución Ji-cuadrada con $n-1$ grados de libertad.

Lin y Sanford (1983) estudiaron la potencia y el nivel de significancia de varias pruebas estadísticas cuando los datos provenían de las distribuciones Poisson, Binomial Negativa o Geométrica. Las pruebas estudiadas fueron: la prueba de la razón de verosimilitudes asumiendo una distribución Poisson, prueba de suma de rangos Kruskal-Wallis y prueba de razón-F. Además, desarrollaron la prueba de la razón de verosimilitudes asumiendo la distribución Binomial Negativa cuando los datos provenían de tal distribución. Cabe mencionar que todas las pruebas desarrolladas son pruebas de bondad de ajuste.

Nakamura y Pérez-Abreu (1993) propusieron un método estadístico gráfico como una guía para elegir la mejor distribución para el ajuste de los datos entre las distribuciones de conteo tales como Poisson, Binomial o Binomial Negativa. Este método esta basado en la forma gráfica del logaritmo de la función generadora de probabilidades empírica, que se define como:

$$\hat{Y}(t) = \log \hat{\phi}(t) = \log \left(\frac{1}{n} \sum_{i=0}^n t^{x_i} \right), \quad 0 < t \leq 1.$$

Las ventajas del método propuesto son las siguientes: no necesita datos agrupados, es una técnica no paramétrica y trabaja de manera razonable con muestras pequeñas. Este método también tiene las siguientes desventajas:

- Es un método subjetivo porque depende de la apreciación de cada persona para elegir entre las distribuciones Poisson, Binomial o Binomial Negativa, a través de observar la forma que tiene la gráfica de $\hat{Y}(t)$. Si la forma de la gráfica es recta, cóncava o convexa entonces elegir a la distribución Poisson, Binomial o Binomial Negativa, respectivamente.
- El intervalo de valores posibles para la variable t está muy limitado, ya que si los datos tienen valores grandes, entonces el intervalo $(0,1]$ deberá ser reducido.

Spinelli y Stephens (1997) propusieron una prueba de bondad de ajuste basadas en las estadísticas de Cramér-Von Mises con la finalidad de probar el siguiente juego de hipótesis: H_0 : la muestra aleatoria proviene de una distribución Poisson con media desconocida. vs H_a : la muestra aleatoria no proviene de una distribución Poisson. Estos autores mostraron que esta prueba empleando las estadísticas mencionadas resulto ser mas potente comparadas con las otras pruebas de bondad de ajuste comunes, como por ejemplo: el Índice de Dispersión de Poisson, las estadísticas basada en la función generadora de probabilidades empírica, la estadística de Pearson, entre otras.

1.2.- PLANTEAMIENTO DEL PROBLEMA

Debido a que se han encontrado algunas dificultades en aplicar las pruebas mencionadas en la sección anterior, surge la iniciativa de construir una prueba de bondad de ajuste que permita afirmar con cierta seguridad que los datos en cuestión se distribuyen de acuerdo al modelo Poisson. De esta manera para la realización de esta investigación, se trabajó con el siguiente Índice de Dispersión:

$$I = \frac{\mu}{\sigma^2}, \text{ donde: } \mu = \text{media poblacional}, \quad \sigma^2 = \text{varianza poblacional} \quad \text{e} \quad \hat{I} = \frac{\bar{X}}{S^2}$$

$$\text{donde: } \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{y} \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Dicho índice fue considerado como la estadística de prueba,
ya que:

- Para el modelo Poisson con $\mu = \sigma^2 = \lambda$ se tiene que:

$$I = 1$$

- Para el modelo Binomial con $\mu = np$ y $\sigma^2 = np(1-p)$ se tiene que:

$$I = \frac{1}{1-p} > 1, p \in (0,1)$$

- Para el modelo Binomial Negativa con $\mu = \frac{r(1-p)}{p}$ y $\sigma^2 = \frac{r(1-p)}{p^2}$ se

tiene que: $I = p < 1$, $p \in (0,1)$ y $r > 0$

En la sección 2.1 (distribuciones de variables de conteo) que se presenta mas adelante se describen brevemente las distribuciones correspondientes a los modelos mencionados.

1.3.- OBJETIVOS

- Proponer una prueba de bondad de ajuste para la distribución Poisson basada en el Índice de Dispersión de Fisher.

- Comparar la potencia de la prueba propuesta con la prueba de bondad de ajuste del Índice de Dispersión de Fisher.

2.- MARCO TEORICO

2.1.- DISTRIBUCIONES DE VARIABLES DE CONTEO

Como ya se mencionó en secciones anteriores las distribuciones Poisson, Binomial y Binomial Negativa son útiles para modelar fenómenos aleatorios en donde las variables aleatorias se refieren a conteos, por lo que a continuación se exponen algunas de sus propiedades.

2.1.1.- DISTRIBUCIÓN POISSON

Una variable aleatoria X que tiene una función de densidad de probabilidades (fdp) de la forma:

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, \dots, \quad \text{donde } \lambda > 0, \text{ tiene una distribución Poisson, con}$$

parámetro λ .

La función generadora de momentos (fgm) de una distribución Poisson, es definida como:

$$m_x(t) = E(e^{tX}) = \sum_x e^{tx} f(x) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t \lambda)^x}{x!} = e^{-\lambda} e^{e^t \lambda} = e^{\lambda(e^t - 1)}$$

para todos los valores reales de t . (El resultado se obtiene de la expresión: $\sum_{x=0}^{\infty} \frac{a^x}{x!} = e^a$)

A partir de la fgm se puede obtener la media μ y la varianza σ^2 de la distribución, esto es:

$E(X) = \mu = m'_x(0)$ y $Var(X) = \sigma^2 = m''_x(0) - \mu^2$, donde $m'_x(0)$ y $m''_x(0)$ son respectivamente la primera y segunda derivadas de la fgm respecto de t y evaluadas en $t = 0$.

De esta forma al realizar lo anterior se puede mostrar que:

$E(X) = Var(X) = \lambda$, es decir, μ y σ^2 son iguales.

2.1.2.- DISTRIBUCIÓN BINOMIAL

Una variable aleatoria X que tiene una fdp de la forma:

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n, \text{ donde } n \text{ es un entero positivo y } 0 < p < 1,$$

tiene una distribución Binomial, con parámetros n y p .

La fgm de una distribución Binomial es definida como:

$$m_x(t) = \sum_{x=0}^n e^{tx} \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=0}^n \binom{n}{x} (pe^t)^x (1-p)^{n-x} = [(1-p) + pe^t]^n$$

para todos los valores reales de t . (El resultado se obtiene de la expresión:

$$\sum_{j=0}^n \binom{n}{j} a^j b^{n-j} = (a+b)^n$$

Derivando la fgm respecto de t y evaluando en $t = 0$, se obtiene μ y σ^2 , de tal manera que:

$$E(X) = np \quad \text{y} \quad \text{Var}(X) = np(1-p)$$

2.1.3.- DISTRIBUCIÓN BINOMIAL NEGATIVA

Una distribución con una fdp de la forma:

$$f(x) = \binom{x+r-1}{x} p^r (1-p)^x, \quad x = 0, 1, \dots, \quad , \text{ donde } r \text{ es un entero positivo fijo, que}$$

representa el número de éxitos, y X es la variable aleatoria que representa el número de no éxitos, es llamada la distribución Binomial Negativa.

La fgm de esta distribución es:

$$m_x(t) = \sum_{x=0}^{\infty} e^{tx} \binom{x+r-1}{x} p^r (1-p)^x = p^r \sum_{x=0}^{\infty} \binom{r+x-1}{x} [(1-p)e^t]^x = p^r [1 - (1-p)e^t]^{-r} = \left[\frac{p}{1 - qe^t} \right]^r$$

donde $q = 1 - p$. (Este resultado se deriva de la expresión:

$$\sum_{j=0}^{\infty} \binom{-n}{j} (-x)^j = \sum_{j=0}^{\infty} \binom{n+j-1}{j} x^j = (1-x)^{-n}, \quad \text{para } -1 < x < 1)$$

Al derivar la fgm respecto de t y evaluando en $t = 0$, se obtiene μ y σ^2 , de tal forma que:

$$\mu = E(X) = \frac{rq}{p} \quad \text{y} \quad \sigma^2 = \text{Var}(X) = \frac{rq}{p^2}$$

2.2.- PRUEBA DE BONDAD DE AJUSTE

2.2.1.- PROBLEMA

Supongamos que para cierto fenómeno aleatorio de conteo, X , se desconoce $f(x)$, entonces surge la pregunta ¿ $f(x)$ es una densidad de Poisson?. Para contestar a la pregunta, se tienen dos posibilidades: se consulta a un adivino para que diga si $f(x)$ es ó no es de Poisson ó se pide permiso a la naturaleza para observar algunos ejemplos de realizaciones de la variable aleatoria X , es decir, se observa experimentalmente el fenómeno algunas veces, y con base en la información obtenida, se determina si los datos contradicen el que $f(x)$ sea una densidad de Poisson. La implementación de la segunda solución, es un tema que en estadística matemática recibe el nombre de prueba de bondad de ajuste. El problema de bondad de ajuste es entonces: con base en una muestra aleatoria X_1, X_2, \dots, X_n , determinar si contradice que $f(x)$ sea Poisson.

En general una prueba de bondad de ajuste es una forma objetiva de evaluar si un conjunto de datos puede ser modelado por una distribución conocida

2.2.2.- DEFINICIÓN

Sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n , asociada con alguna función de distribución desconocida, denotada por $F(x)$.

Sea $F^*(x)$ la distribución conocida, se desea probar el siguiente juego de hipótesis,

$$H_0: F(x) = F^*(x)$$

$$H_a: F(x) \neq F^*(x).$$

Sea $T = t(x)$ una estadística de prueba.

Para el caso en que la distribución de la estadística de prueba es de una cola Rechazar H_0 si $T < K\alpha$ y no rechazar H_0 si $T \geq K\alpha$, donde: $K\alpha$ es una constante crítica que se obtiene de tal forma que la prueba sea de tamaño α , es decir, $P(T < K\alpha \mid H_0 \text{ es verdadera}) \leq \alpha$, para cualquier distribución $F^*(x)$.

Para el caso en que la distribución de la estadística de prueba es de dos colas entonces rechazamos H_0 si $T < K_{\alpha/2}$ o si $T > K_{1-\alpha/2}$ y no rechazamos H_0 si $K_{\alpha/2} \leq T \leq K_{1-\alpha/2}$, donde: K_{α} es una constante crítica que se obtiene de tal forma que la prueba sea de tamaño α , es decir, $P(T < K_{\alpha/2} \text{ o } T > K_{1-\alpha/2} \mid H_0 \text{ es verdadera}) = P(T < K_{\alpha/2} \mid H_0 \text{ es verdadera}) + P(T > K_{1-\alpha/2} \mid H_0 \text{ es verdadera}) \leq \alpha$.

Recuerde que en una prueba de bondad de ajuste no se tiene especificada la hipótesis alternativa.

2.3.- ALGUNAS DEFINICIONES RELACIONADAS CON LAS PRUEBAS DE HIPÓTESIS

Una Prueba de Hipótesis para probar H_0 vs H_a esta dada por una partición de X en X_r y X_a , donde: $X = \{\underline{x} : \underline{x} \text{ es una realización de } X\}$, \underline{x} es un vector de observaciones con densidad $f(\underline{x}, \theta)$, $\theta \in \Omega = (a, b) \subset \mathfrak{R}$, a y b son valores conocidos que pertenecen a los números reales, X_r es la región de rechazo y X_a es la región de no rechazo.

Una Regla de Decisión esta dada por: rechazar H_0 si $\underline{x} \in X_r$ y no rechazar H_0 si $\underline{x} \in X_a$.

Generalmente cuando usamos una prueba de hipótesis existe una probabilidad de cometer un error. A saber estamos sujetos a cometer dos tipos de errores: Error Tipo I : Rechazar H_0 cuando en realidad es verdadera y Error Tipo II : No rechazar H_0 cuando en realidad es falsa. De esta manera la Probabilidad de cometer el Error Tipo I se expresa como: $P(\text{Error Tipo I}) = \alpha$ y la Probabilidad de cometer el Error Tipo II como: $P(\text{Error Tipo II}) = \beta$. Por otra parte es razonable identificar una prueba que minimice las probabilidades de ambos errores respecto a todas las pruebas posibles, es decir se desea encontrar una prueba de hipótesis tal que tanto α como β sean mínimas, sin embargo, esto ocurre muy pocas veces debido a que si se disminuye una de las probabilidades la otra probabilidad aumenta, por esa razón las pruebas de hipótesis se construyen de un tamaño dado.

La Función Potencia $\varphi(\theta)$, donde θ es el vector de parámetros que se desea inferir, se define como $\varphi(\theta) = P(\text{Rechazar } H_0 / \theta)$

Una prueba es de tamaño α si la función de potencia evaluada cuando la hipótesis nula es verdadera, es menor o igual a α .

2.4.- EL MÉTODO DELTA PARA CALCULAR DISTRIBUCIONES ASINTÓTICAS

El método delta es una técnica general importante para calcular la distribución asintótica de una función de variables aleatorias, las cuales provienen de una distribución en específico.

2.4.1.- CASO UNIVARIADO

Sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n independientes e idénticamente distribuidas, entonces por el teorema central del límite

$$\bar{X} \approx N\left(\mu, \frac{\sigma^2}{n}\right) \text{ cuando } n \rightarrow \infty.$$

Además sea $g(X) = g(a) + (X - a)g'(a) + \dots$ la expansión de la serie de Taylor de $g(X)$, entonces $g(\bar{X}) = g(\mu) + (\bar{X} - \mu)g'(\mu) + \dots$ es la expansión de la serie de Taylor de $g(\bar{X})$.

Por lo tanto por el teorema del método delta para el caso univariado se tiene que:

$$g(\bar{X}) \approx N\left(g(\mu), (g'(\mu))^2 \frac{\sigma^2}{n}\right),$$

$$\text{donde: } g(\mu) = E(g(x)) \quad \text{y} \quad g'(\mu) = E(g'(x)) = E\left(\frac{\partial}{\partial x} g(x)\right)$$

2.4.2.- CASO BIVARIADO

Sea $\begin{pmatrix} X_1 \\ Y_1 \end{pmatrix}, \begin{pmatrix} X_2 \\ Y_2 \end{pmatrix}, \dots, \begin{pmatrix} X_n \\ Y_n \end{pmatrix}$ una muestra aleatoria bivariada de tamaño n

independientes e idénticamente distribuidas, entonces por el teorema central del límite

$$\begin{pmatrix} \bar{X} \\ \bar{Y} \end{pmatrix} \approx \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \frac{\Sigma}{n} \right) \text{ cuando } n \rightarrow \infty.$$

Además sea $g(X, Y) = g(a, b) + (X - a)g_{10}(a, b) + (Y - b)g_{01}(a, b) + \dots$ la expansión de la serie de Taylor de $g(X, Y)$, entonces $g(\bar{X}, \bar{Y}) = g(\mu_1, \mu_2) + (\bar{X} - \mu_1)g_{10}(\mu_1, \mu_2) + (\bar{Y} - \mu_2)g_{01}(\mu_1, \mu_2) + \dots$ es la expansión de la serie de Taylor de $g(\bar{X}, \bar{Y})$.

Por lo tanto por el teorema de la teoría del método delta para el caso bivariado se tiene que $g(\bar{X}, \bar{Y}) \approx N\left(g(\mu_1, \mu_2), (g_{10}(\mu_1, \mu_2), g_{01}(\mu_1, \mu_2)) \frac{\sum}{n} (g_{10}(\mu_1, \mu_2), g_{01}(\mu_1, \mu_2))'\right)$,

donde: $g(\mu_1, \mu_2) = E(g(x, y))$, $g_{10}(\mu_1, \mu_2) = E(g_{10}(x, y)) = E\left(\frac{\partial}{\partial x} g(x, y)\right)$

y $g_{01}(\mu_1, \mu_2) = E(g_{01}(x, y)) = E\left(\frac{\partial}{\partial y} g(x, y)\right)$

3.- PRUEBA DE BONDAD DE AJUSTE DEL ÍNDICE DE DISPERSIÓN I

3.1.- DISTRIBUCIÓN DE LA ESTADÍSTICA DE PRUEBA

Considerando la definición expuesta en la sección 2.2.2 tenemos que la prueba requiere de la constante crítica $K\alpha$, es decir de los percentiles de la distribución I bajo la hipótesis nula. Para la obtención de estos es necesario primero obtener la distribución de la estadística de prueba ya sea analíticamente o empíricamente. En esta investigación se revisaran las dos formas:

3.1.1.- OBTENCIÓN EN FORMA ANALÍTICA

Tenemos que la estadística de prueba es $I = \frac{\bar{X}}{S^2}$ donde: $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ y

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \left(\frac{1}{n} \sum_{i=1}^n X_i \right)^2 \quad \text{cuando } n \rightarrow \infty,$$

entonces $I = \frac{\bar{X}}{Y - \bar{X}^2}$ donde $Y = \frac{1}{n} \sum_{i=1}^n X_i^2$.

Ahora considerando lo expuesto en la sección 2.4.2.- tenemos que

$$g(\bar{X}, Y) = I = \frac{\bar{X}}{Y - \bar{X}^2} \quad \text{y} \quad g(X, Y) = \frac{X}{Y - X^2}, \quad \text{entonces}$$

$$g_{10}(X, Y) = \frac{\partial g(X, Y)}{\partial X} = \frac{Y + X^2}{(Y - X^2)^2} \quad \text{y} \quad g_{01}(X, Y) = \frac{\partial g(X, Y)}{\partial Y} = -\frac{X}{(Y - X^2)^2}.$$

Por otro lado sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n independientes e idénticamente distribuidas $\text{Poisson}(\lambda)$, entonces $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \approx N\left(\lambda, \frac{\lambda}{n}\right)$ y

$$Y = \frac{1}{n} \sum_{i=1}^n X_i^2 \approx N\left(\lambda + \lambda^2, \frac{\lambda + 6\lambda^2 + 4\lambda^3}{n}\right), \quad \text{ya que:}$$

$$E(X) = \lambda, \quad E(X^2) = \lambda + \lambda^2, \quad E(X^3) = \lambda + 3\lambda^2 + \lambda^3 \quad \text{y} \quad E(X^4) = \lambda + 7\lambda^2 + 6\lambda^3 + \lambda^4$$

De esta manera tenemos que:

$$g(\mu_1, \mu_2) = \frac{\lambda}{\lambda + \lambda^2 - \lambda^2} = 1 \text{ puesto que } g(\bar{X}, \bar{Y}) = \frac{\bar{X}}{\bar{Y} - \bar{X}^2},$$

$$g_{10}(\mu_1, \mu_2) = \frac{\lambda + \lambda^2 + \lambda^2}{(\lambda + \lambda^2 - \lambda^2)^2} = \frac{1 + 2\lambda}{\lambda} \text{ ya que } g_{10}(\bar{X}, \bar{Y}) = \frac{\bar{Y} + \bar{X}^2}{(\bar{Y} - \bar{X}^2)^2} \text{ y}$$

$$g_{01}(\mu_1, \mu_2) = -\frac{\lambda}{(\lambda + \lambda^2 - \lambda^2)^2} = -\frac{1}{\lambda} \text{ pues } g_{01}(\bar{X}, \bar{Y}) = -\frac{\bar{X}}{(\bar{Y} - \bar{X}^2)^2}.$$

También tenemos que la matriz de varianzas y covarianzas es:

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{21} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} = \begin{pmatrix} \lambda & \lambda + 2\lambda^2 \\ \lambda + \lambda^2 & \lambda + 6\lambda^2 + 4\lambda^3 \end{pmatrix} \text{ ya que:}$$

$$\sigma_{11} = E(X), \quad \sigma_{21} = \sigma_{12} = E(X^3) - E(X^2)E(X) \text{ y } \sigma_{22} = E(X^4) - (E(X^2))^2.$$

De esta forma después de realizar un poco de álgebra encontramos que:

$$\begin{pmatrix} g_{10}(\mu_1, \mu_2) & g_{01}(\mu_1, \mu_2) \end{pmatrix} \Sigma \begin{pmatrix} g_{10}(\mu_1, \mu_2) & g_{01}(\mu_1, \mu_2) \end{pmatrix}' =$$

$$\begin{pmatrix} \frac{1+2\lambda}{\lambda} & -\frac{1}{\lambda} \end{pmatrix} \begin{pmatrix} \lambda & \lambda + 2\lambda^2 \\ \lambda + \lambda^2 & \lambda + 6\lambda^2 + 4\lambda^3 \end{pmatrix} \begin{pmatrix} \frac{1+2\lambda}{\lambda} \\ \frac{1}{\lambda} \\ -\frac{1}{\lambda} \end{pmatrix} = 2$$

Finalmente considerando el resultado de que:

$$g(\bar{X}, \bar{Y}) \approx N\left(g(\mu_1, \mu_2), (g_{10}(\mu_1, \mu_2), g_{01}(\mu_1, \mu_2)) \frac{\Sigma}{n} (g_{10}(\mu_1, \mu_2), g_{01}(\mu_1, \mu_2))'\right), \text{ concluimos}$$

$$\text{que } g(\bar{X}, \bar{Y}) = I \approx N\left(1, \frac{2}{n}\right).$$

Por lo tanto la estadística de prueba I bajo la hipótesis nula se distribuye asintóticamente normal con media 1 y varianza $\frac{2}{n}$, de donde deducimos que no depende del parámetro λ .

De esta forma utilizando el lenguaje de programación S_PLUS 2000 y mediante el Programa 1 que se encuentra en la sección 9.2 (anexo B) se obtuvieron los cuantiles de la distribución $N\left(1, \frac{2}{n}\right)$ para los siguientes tamaños de muestra 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 300, 500, 1000, 5000 y 10000. Los resultados se encuentran en la Gráfica 4 y Tabla 8 de la sección 9.1 (anexo A). Esto se realizó con el propósito de compararlos con los obtenidos de la distribución de la estadística de prueba en forma empírica. Dicha comparación se presenta más adelante.

3.1.2.- OBTENCIÓN EN FORMA EMPIRICA

Para este caso los percentiles fueron estimados mediante simulación por el Método Monte Carlo. La simulación se hizo en el lenguaje de programación S_PLUS 2000 mediante el Programa 2 que se encuentra en la sección 9.2 (anexo B). Para lograr dicha estimación se generaron muestras de tamaño 10 , 20, 30, 40, 50,60,70, 80,90, 100, 150, 200, 300, 500, 1000, 5000 y 10000; de la distribución Poisson. Para cada tamaño de muestra se obtuvo el valor de I en 20,000 ocasiones y con esta información se obtuvo la distribución empírica de I, con la cual se estimaron los cuantiles 0.0125, 0.025, 0.05, 0.1, 0.2, 0.4, 0.5, 0.6, 0.8, 0.9, 0.95, 0.975 y 0.9875. Para lo anterior se empleó el siguiente Algoritmo (1):

- 1.- Generar una muestra aleatoria de la distribución Poisson (λ).
- 2.- Obtener el valor I.
- 3.- Repetir los dos pasos anteriores, 20,000 veces.
- 4.- Con la información del paso anterior, obtener los cuantiles de la distribución empírica del valor I

Cabe mencionar que se consideraron los siguientes tres valores de λ : $\lambda=1$, $\lambda=2$ y $\lambda=3$. De esta forma se tiene que los histogramas y cuantiles de la distribución de la estadística de prueba I bajo H_0 (para los diferentes valores de λ) son los que se presentan en las Gráficas 5, 6 y 7 y Tablas 9, 10 y 11 que se muestran en la sección 9.1 (anexo A).

De los resultados obtenidos se puede observar que la diferencia entre los cuantiles para los diferentes valores de λ en cada tamaño de muestra es mínima por ejemplo para $n=60$ y $\alpha=0.025$ tenemos que los cuantiles para $\lambda=1$, $\lambda=2$ y $\lambda=3$ respectivamente son: 0.7084, 0.7218 y 0.7159. De esta manera como se demostró en la sección anterior se reafirma que efectivamente la distribución de la estadística de prueba bajo la hipótesis nula no depende del parámetro λ .

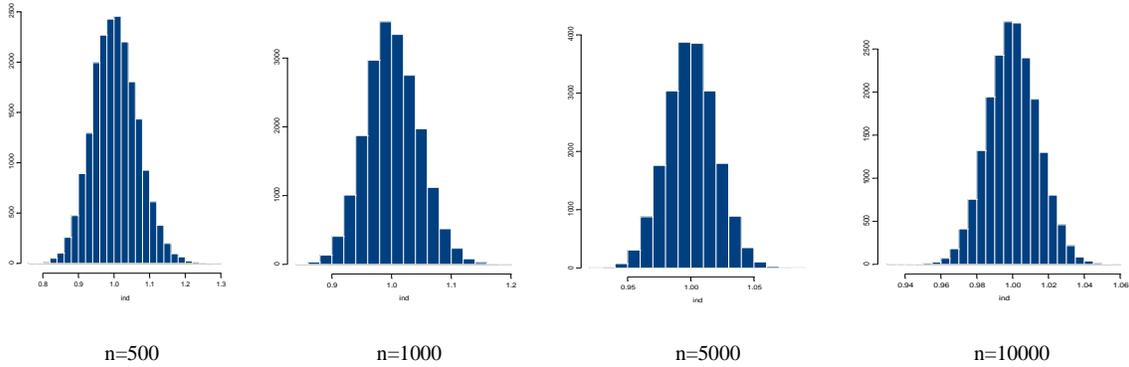
También cabe mencionar que conforme aumentamos el tamaño de la muestra en ambos valores de λ la distribución de la estadística de prueba tiende a una distribución normal con media 1 y varianza $\frac{2}{n}$ ya que si comparamos estos cuantiles con los obtenidos directamente de la distribución $N\left(1, \frac{2}{n}\right)$ podemos notar que prácticamente son los mismos como se puede observar a manera de ejemplo en la siguiente Tabla 1 y Figura 1.

Tabla 1.- Comparación de los cuantiles de la distribución de la estadística de prueba I^* se distribuye $Poisson(\lambda=1)$, estimados con 20,000 muestras, por el método Monte Carlo vs

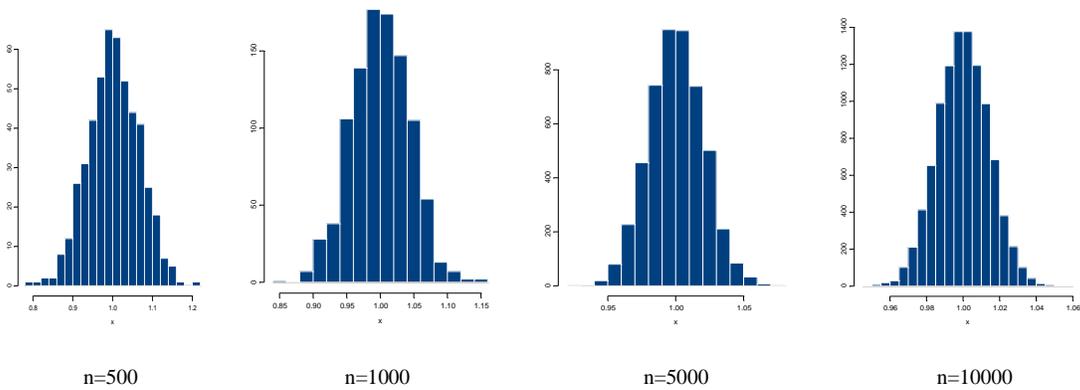
los cuantiles de la distribución $N\left(1, \frac{2}{n}\right)$.

n/α	0.0125	0.025	0.05	0.1	0.2	0.4	0.5
SIMULACION							
500	0.8686	0.8842	0.9029	0.9235	0.9494	0.9861	1.0021
1000	0.9061	0.9173	0.9308	0.9455	0.9641	0.9891	1.0006
5000	0.9564	0.9617	0.9675	0.9748	0.9836	0.9951	1.0002
10000	0.9690	0.9729	0.9774	0.9822	0.9883	0.9967	1.0001
EXACTA							
500	0.8673	0.8797	0.8959	0.9190	0.9493	0.9858	1.0022
1000	0.9039	0.9134	0.9311	0.9469	0.9631	0.9879	1.0004
5000	0.9562	0.9619	0.9672	0.9743	0.9831	0.9950	1.0002
10000	0.9690	0.9724	0.9768	0.9820	0.9879	0.9965	1.0000
n/α	0.5	0.6	0.8	0.9	0.95	0.975	0.9875
SIMULACION							
500	1.0021	1.0184	1.0566	1.0874	1.1121	1.1330	1.1534
1000	1.0006	1.0120	1.0393	1.0598	1.0765	1.0925	1.1065
5000	1.0002	1.0052	1.0167	1.0255	1.0331	1.0397	1.0453
10000	1.0001	1.0037	1.0120	1.0183	1.0237	1.0281	1.0320
EXACTA							
500	1.0022	1.0182	1.0582	1.0868	1.1080	1.1249	1.1431
1000	1.0004	1.0122	1.0372	1.0538	1.0689	1.0788	1.0965
5000	1.0002	1.0052	1.0175	1.0262	1.0331	1.0398	1.0459
10000	1.0000	1.0037	1.0119	1.0181	1.0232	1.0273	1.0313

SIMULACION



EXACTA



Gráfica 1.- Comparación de los histogramas de la estadística de prueba I^* bajo H_0 de que x se distribuye Poisson ($\lambda=1$), estimados con 20,000 muestras, por el método Monte Carlo vs los histogramas de la distribución $N\left(1, \frac{2}{n}\right)$.

La comparación para los otros tamaños de muestra y valores de λ la podemos realizar al contrastar la Tabla 8 con las Tablas 9, 10 y 11 que se muestran en la sección 9.1 (anexo A).

3.2.- CONSTRUCCIÓN DE LA PRUEBA DE BONDAD DE AJUSTE

Sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n , asociada con alguna función de distribución desconocida, denotada por $F(x)$.

Sea $F^*(x)$ la distribución Poisson, se desea probar el siguiente juego de hipótesis,

$H_0: F(x) = F^*(x)$

$H_a: F(x) \neq F^*(x)$.

Sea $I = \frac{\bar{X}}{S^2} = T = t(x)$ una estadística de prueba.

Como se puede notar en la sección anterior la distribución de la estadística de prueba es de dos colas entonces rechazamos H_0 si $I < K_{\alpha/2}$ o si $I > K_{1-\alpha/2}$ y no rechazamos H_0 si $K_{\alpha/2} \leq I \leq K_{1-\alpha/2}$, donde: K_{α} es una constante crítica que se obtiene de tal forma que la prueba sea de tamaño α , es decir, $P(I < K_{\alpha/2} \text{ o } I > K_{1-\alpha/2} \mid H_0 \text{ es verdadera}) = P(I < K_{\alpha/2} \mid H_0 \text{ es verdadera}) + P(I > K_{1-\alpha/2} \mid H_0 \text{ es verdadera}) \leq \alpha$.

Note que la prueba requiere de la constante crítica $K\alpha$, es decir de los percentiles de la distribución I bajo la hipótesis nula. La estimación de estos se dan en la Tabla 8 en forma exacta y en las Tablas 9, 10 y 11 por simulación, dichas tablas se encuentran en la sección 9.1 (anexo A). De esta manera si utilizamos dichas tablas para encontrar $K\alpha$ sería un poco engorroso. Por lo que se construyó una prueba en donde se utilizan los valores de la función de distribución normal estándar. Dicha prueba se presenta a continuación.

Como se demostró en la sección 3.1.1 tenemos que $g(\bar{X}, \bar{Y}) = \frac{\bar{X}}{\bar{Y} - \bar{X}^2} = \frac{\bar{X}}{S^2} = I \approx N\left(1, \frac{2}{n}\right)$ entonces estandarizando tenemos que

$$\frac{\frac{\bar{X}}{\bar{Y} - \bar{X}^2} - 1}{\sqrt{\frac{2}{n}}} \approx N(0,1) \quad \Rightarrow \quad \sqrt{\frac{n}{2}} \left(\frac{\bar{X} - (\bar{Y} - \bar{X}^2)}{\bar{Y} - \bar{X}^2} \right) = \sqrt{\frac{n}{2}} \left(\frac{\bar{X} - S^2}{S^2} \right) = \sqrt{\frac{n}{2}} \left(\frac{\bar{X}}{S^2} - 1 \right) \approx N(0,1)$$

De esta forma sea $I^* = Z_{cal} = \sqrt{\frac{n}{2}} \left(\frac{\bar{X}}{S^2} - 1 \right) = T = t(x)$ una estadística de prueba y $Z_{tab} = Z\alpha/2$ un valor de las tablas de la normal estándar, entonces rechazamos H_0 si $I^* < Z\alpha/2$ o si $I^* > Z1-\alpha/2$ y no rechazamos H_0 si $Z\alpha/2 \leq I^* \leq Z1-\alpha/2$, donde: $Z\alpha$ es una constante crítica que se obtiene de tal forma que la prueba sea de tamaño α , es decir, $P(I^* < Z\alpha/2 \text{ o } I^* > Z1-\alpha/2 \mid H_0 \text{ es verdadera}) = P(I^* < Z\alpha/2 \mid H_0 \text{ es verdadera}) + P(I^* > Z1-\alpha/2 \mid H_0 \text{ es verdadera}) \leq \alpha$.

Por ejemplo para $\alpha = 0.05$ tenemos que:

a).- $P(I^* < -1.96) = 0.025$

b).- $P(I^* > 1.96) = 0.025$

c).- $P(I^* < -1.96) + P(I^* > 1.96) = 0.05$

d).- $P(-1.96 \leq I^* \leq 1.96) = 0.95$

e).- $P(I^* < -1.96) + P(I^* > 1.96) + P(-1.96 \leq I^* \leq 1.96) = 1.00$

3.3.- TAMAÑO DE LA PRUEBA

Una parte importante de una prueba de bondad de ajuste es saber cual es el tamaño de la prueba. Para esto se estimó la probabilidad de rechazar la hipótesis nula cuando las muestras aleatorias pertenecen a una distribución Poisson (λ), pues como se mencionó en la sección 2.3 (algunas definiciones relacionadas con las pruebas de hipótesis), una prueba es de tamaño α si la función de potencia evaluada cuando la hipótesis nula es verdadera, es menor o igual a α . Para lo anterior se generaron muestras de tamaño 30, 50, 80, 100, 300 y 500 las cuales se sometieron a la prueba para los niveles de significancia 0.025, 0.05, 0.075, 0.1 y 0.125, mediante el siguiente Algoritmo(2):

-
-
- 1.- Generar una muestra aleatoria de la distribución Poisson (λ).
 - 2.- Obtener el valor I^* .
 - 3.- Repetir los dos pasos anteriores, 10,000 veces.
 - 4.- Para cada una muestra de las 10,000 obtenidas, se compara el valor de I^* con $Z\alpha/2$ de

tal manera que:

Si $I^* < Z\alpha/2$ o si $I^* > Z1-\alpha/2$, entonces la muestra toma el valor 1.

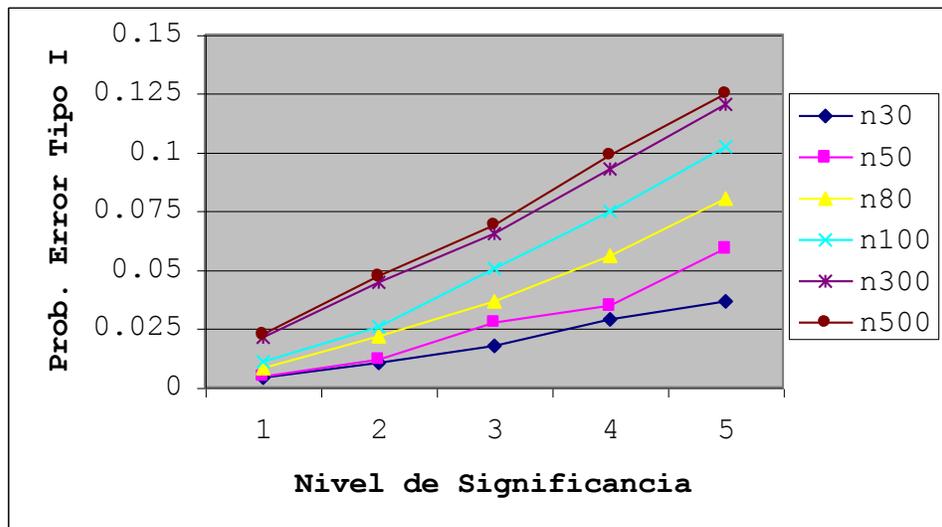
Si $Z\alpha/2 \leq I^* \leq Z1-\alpha/2$, entonces la muestra toma el valor 0.

- 5.- Una vez realizada la condición del paso anterior sumar los valores igual a 1.
- 6.- Dividir el total de valores igual a 1 entre 10,000 y el resultado es la probabilidad estimada del error tipo 1.

Para la obtención de estos resultados, la simulación se hizo en el lenguaje de programación S_PLUS 2000 mediante el Programa 3 que se encuentra en la sección 9.2 (anexo B) para diferentes valores de λ , dichos resultados se presentan en la sección 9.1 (anexo A) y uno a manera de ejemplo en la siguiente Tabla 2 y Grafica 2:

Tabla 2.- Ejemplo del tamaño de la prueba para la distribución Poisson($\lambda = 0.2$)

	ns	ns1	ns2	ns3	ns4	ns5
N	PET1	0.025	0.05	0.075	0.1	0.125
30		0.0039	0.0105	0.0177	0.0287	0.0364
50		0.0046	0.0117	0.0275	0.0348	0.0588
80		0.0079	0.0215	0.0366	0.0559	0.0802
100		0.0106	0.0257	0.0504	0.0749	0.1021
300		0.0213	0.0446	0.0654	0.0927	0.1202
500		0.0223	0.0472	0.0690	0.0987	0.1248



Gráfica 2.- Ejemplo probabilidad estimada del error tipo I de la prueba cuando la muestra proviene de la distribución Poisson($\lambda = 0.2$)

Se puede notar que para todos los tamaños de muestra la probabilidad estimada del error tipo I cuando los datos pertenecen a la distribución Poisson(0.2) es menor al nivel de significancia señalado.

Al igual que en el ejemplo mostrado, para valores de λ menores que 1, la probabilidad estimada del error tipo I cuando los datos pertenecen a la distribución Poisson es menor al nivel de significancia señalado. Por lo tanto como la probabilidad estimada del error tipo I es menor al nivel de significancia (α), en estos valores del parámetro de la distribución Poisson podemos decir que existe evidencia de que tenemos una prueba de tamaño α .

3.4.- POTENCIA DE LA PRUEBA

Otra parte importante de una prueba de bondad de ajuste es saber cual es la potencia de la prueba, dicha potencia solo es posible obtenerla si se define la hipótesis alternativa.

De esta forma como la estadística de prueba $I = 1$ para el modelo Poisson, $I > 1$ para el modelo Binomial y $I < 1$ para el modelo Binomial Negativa, sólo se generaron muestras aleatorias de estas distribuciones, con el propósito de estimar la probabilidad de rechazar la hipótesis nula y así observar el comportamiento de la prueba. Para esto se generaron muestras de tamaño 10, 20, 30, 40, 50 y 80 las cuales se sometieron a la prueba para los niveles de significancia 0.025, 0.05, 0.075, 0.1 y 0.125. Para lo anterior se utilizó el siguiente Algoritmo (3):

-
-
- 1.- Generar una muestra aleatoria diferente a la distribución Poisson.
 - 2.- Obtener el valor I^* .
 - 3.- Repetir los dos pasos anteriores, 10,000 veces.
 - 4.- Para cada una muestra de las 10,000 obtenidas, se compara el valor de I^* con $Z\alpha/2$ de tal manera que:

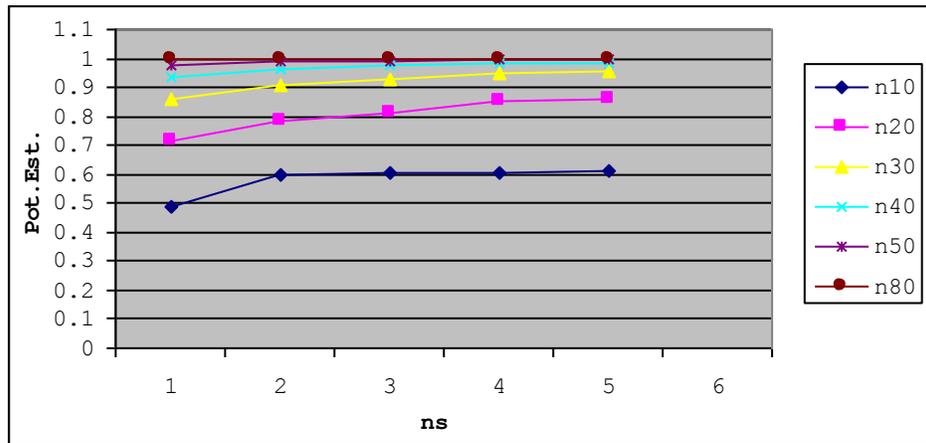
Si $I^* < Z\alpha/2$ o si $I^* > Z1-\alpha/2$, entonces la muestra toma el valor 1.

Si $Z\alpha/2 \leq I^* \leq Z1-\alpha/2$, entonces la muestra toma el valor 0.
 - 5.- Una vez realizada la condición del paso anterior, sumar los valores igual a 1.
 - 6.- Dividir el total de valores igual a 1 entre 10,000 y el resultado es la potencia estimada de la prueba.

Para la obtención de los resultados solo se consideraron las distribuciones Binomial y Binomial Negativa para diferentes parámetros, ya que como se mencionó en la sección 1.2 (planteamiento del problema), para el modelo Poisson $I = 1$, para el modelo Binomial $I > 1$ y para el modelo Binomial Negativa $I < 1$, estos se generaron en el lenguaje de programación S_PLUS 2000 mediante los Programas 4 y 5 que se encuentra en la sección 9.2 (anexo B), dichos resultados se presentan en la sección 9.1 (anexo A) y uno a manera de ejemplo en la siguiente Tabla 3 y Grafica 3:

Tabla 3.- Ejemplo de la potencia de la prueba cuando la H_a es una Binomial con media 1 considerando los valores de la Normal Estándar.

N	ns	ns1	ns2	ns3	ns4	ns5
	ot.Est.	0.025	0.05	0.075	0.1	0.125
10		0.4861	0.5976	0.6070	0.6077	0.6133
20		0.7164	0.7860	0.8093	0.8534	0.8621
30		0.8573	0.9067	0.9283	0.9474	0.9564
40		0.9359	0.9604	0.9775	0.9811	0.9865
50		0.9742	0.9875	0.9932	0.9953	0.9968
80		0.9990	0.9999	1.0000	1.0000	1.0000



Gráfica 3.- Ejemplo de la potencia de la prueba cuando la H_a es una Binomial con media 1 considerando los valores de la Normal Estándar.

Note que la potencia de la prueba estimada es mayor de 0.8 para todos los niveles de significancia cuando el tamaño de la muestra es mayor o igual 30.

En general, para los diferentes parámetros considerados de las distribuciones Binomial y Binomial Negativa, la potencia de la prueba estimada en promedio es mayor de 0.5 para todos los niveles de significancia y todos los tamaños de muestra. Además note también que en todos los casos al aumentar el tamaño de la muestra, aumenta la potencia. Por lo tanto podemos decir que la prueba es consistente respecto al tamaño de la muestra

4.- COMPARACIÓN DE LA POTENCIA DE LA PRUEBA PROPUESTA VS PRUEBA DE BONDAD DE AJUSTE DEL ÍNDICE DE DISPERSIÓN DE FISHER

Como se puede notar en la parte de antecedentes existen diferentes pruebas de bondad de ajuste realizadas con la finalidad de contrastar el siguiente juego de hipótesis: H_0 : la prueba aleatoria proviene de una distribución Poisson vs H_a : la prueba aleatoria no proviene de una distribución Poisson. Para cada una de estas esta claro que se tiene una estadística de prueba, entre ellas el Índice de Dispersión de Fisher, la cual se tomo en cuenta para realizar el estudio comparativo de la potencia.

4.1.- PRUEBA DE BONDAD DE AJUSTE DEL ÍNDICE DE DISPERSIÓN DE FISHER

Sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n , asociada con alguna función de distribución desconocida, denotada por $F(x)$.

HIPÓTESIS: Sea $F^*(x)$ la distribución Poisson, se desea probar el siguiente juego

de hipótesis: $H_0: F(x) = F^*(x)$

$H_a: F(x) \neq F^*(x)$.

ESTADÍSTICA DE PRUEBA: Una estadística para probar si los datos provienen de una

distribución Poisson, es el Índice de Dispersión Poisson, debido a Fisher ,

definido por: $R = \frac{nS^2}{\bar{X}}$ donde $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ y $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

REGLA DE DECISIÓN: La hipótesis H_0 es rechazada si $R < R^*$ donde $R^* = \chi^2_{n-1, 1-\alpha}$ es un

Valor de las tablas de la distribución Ji-cuadrada.

De esta manera la potencia se estimo generando muestras aleatorias da la distribución Binomial y Binomial Negativa, con el propósito de estimar la probabilidad de rechazar H_0 . Para esto se generaron muestras de tamaño 10, 20, y 30 las cuales se sometieron a la prueba para los niveles de significancia 0.025, 0.05, y 0.125. Para lo anterior se utilizo el siguiente Algoritmo (4):

- 1.- Generar una muestra aleatoria diferente a la distribución Poisson.
- 2.- Obtener el valor R .
- 3.- Repetir los dos pasos anteriores, 10,000 veces.
- 4.- Para cada una muestra de las 10,000 obtenidas, se compara el valor de R con R^* de tal manera que:

Si $R < R^*$, entonces la muestra toma el valor 1.

Si $R \geq R^*$, entonces la muestra toma el valor 0.

- 5.- Una vez realizada la condición del paso anterior sumar los valores igual a 1.
- 6.- Dividir el total de valores igual a 1 entre 10,000 y el resultado es la potencia estimada de la prueba.

La obtención de los resultados se hizo en el lenguaje de programación S_PLUS 2000 mediante el programa 6 que se encuentra en la sección 9.2 (anexo B), dichos resultados se presentan en las siguientes Tablas:

Tabla 4.- Ejemplo de la potencia de la prueba cuando la H_a es una Binomial con media 1 considerando los valores de la distribución Ji-cuadrada.

	ns	ns1	ns2	ns3
N	Pot.Est.	0.025	0.05	0.1
10		0.0950	0.1453	0.3122
20		0.3105	0.5131	0.7150
30		0.6191	0.7815	0.8939

De esta forma si comparamos los valores de la tabla de la sección anterior con los de esta, como se presenta en la siguiente tabla para el tamaño de muestra igual a 30, se puede notar que estos son menores tanto para los diferentes tamaños de muestra como para los diferentes niveles de significancia.

Tabla 5 .- Comparación de la potencia de la prueba propuesta vs prueba de bondad de ajuste del índice de dispersión de Fisher cuando H_a : Binomial con media 1

Prueba de bondad de ajuste	Índice de dispersión propuesto			Índice de dispersión de Fisher		
	0.025	0.05	0.1	0.025	0.05	0.1
n.s						
H_a : Binomial con media 1	0.8573	0.9067	0.9474	0.6191	0.7815	0.8939

5.- EJEMPLO DE APLICACIÓN DE LA PRUEBA DE BONDAD DE AJUSTE PROPUESTA

Sea X_1, X_2, \dots, X_n una muestra aleatoria de tamaño n , asociada con alguna función de distribución desconocida, denotada por $F(x)$.

HIPÓTESIS: Sea $F^*(x)$ la distribución Poisson, se desea probar el siguiente juego

de hipótesis:

$$H_0: F(x) = F^*(x)$$

$$H_a: F(x) \neq F^*(x).$$

$$\text{ESTADÍSTICA DE PRUEBA: } Z_{\text{cal}} = \sqrt{\frac{n}{2}} \left(\frac{\bar{X}}{S^2} - 1 \right)$$

REGLA DE DECISIÓN: La hipótesis H_0 es rechazada si $Z_{\text{cal}} < Z_{\text{tab}}$ considerando

valores negativos ó si $Z_{\text{cal}} > Z_{\text{tab}}$ considerando valores positivos

donde $Z_{\text{tab}} = Z_{\alpha/2}$ un valor de las tablas de la normal estándar

De esta forma a manera de ejemplo en esta sección se analizarán dos ejemplos para mostrar la aplicación de la prueba de bondad de ajuste propuesta.

EJEMPLO 1.- Los datos del siguiente cuadro representan los números de muertes de los soldados por patadas de caballo en el ejército Prusiano para el periodo 1875-1894; obtenidos de registros de 14 cuerpos de ejército (Andrews y Herzberg), 1985).

Tabla 6.- Números de muertes debidas a patadas de caballos en el ejército Prusiano.

Números de muertes	Frecuencia
0	144
1	91
2	32
3	11
4	2
Total	280

Con la información anterior tenemos lo siguiente:

$$n = 280$$

$$\bar{X} = 0.7$$

$$S^2 = 0.7627$$

Entonces tenemos que:

$$Z_{cal} = \sqrt{\frac{n}{2}} \left(\frac{\bar{X}}{S^2} - 1 \right) = \sqrt{\frac{280}{2}} \left(\frac{0.7}{0.7627} - 1 \right) = -0.9726$$

Además de las tablas de la normal estándar considerando un nivel de significancia de $\alpha = 0.05$ obtenemos que:

$$Z_{tab} = Z_{\alpha / 2} = Z_{0.025} = -1.96$$

Por lo tanto como $Z_{cal} = -0.9726 > Z_{tab} = -1.96$ entonces no se rechaza H_0 , es decir, con un nivel de confiabilidad del 95% concluimos que los datos de muertes de los soldados por patadas de caballos se distribuyen de acuerdo al modelo Poisson.

Cónsul y Jain (1973) verificaron por medio de la prueba de bondad de ajuste de Ji-cuadrada, que los datos del cuadro anterior tienen un buen ajuste al modelo Poisson. Por consiguiente, la conclusión obtenida por la prueba de bondad de ajuste propuesta es consistente con la conclusión de Cónsul y Jain.

EJEMPLO 2.- Los datos del siguiente cuadro representan los números de excedencias del ozono arriba del umbral nivel 16, ocurridos por semestre durante 11 años en la ciudad de Houston, Texas (Smith, 1989).

Tabla 7.- Conteos de excedencias del ozono arriba del umbral nivel 16.

Conteos de excedencias	Frecuencia
2	5
3	9
5	4
6	3
9	1
Total	22

Con la información anterior tenemos lo siguiente:

$$n = 22$$

$$\bar{X} = 3.8182$$

$$S^2 = 3.2987$$

Entonces tenemos que:

$$Z_{cal} = \sqrt{\frac{n}{2}} \left(\frac{\bar{X}}{S^2} - 1 \right) = \sqrt{\frac{22}{2}} \left(\frac{3.8182}{3.2987} - 1 \right) = 0.5223$$

Además de las tablas de la normal estándar considerando un nivel de significancia de $\alpha = 0.05$ obtenemos que:

$$Z_{tab} = Z_{\alpha/2} = Z_{0.025} = 1.96$$

Por lo tanto como $Z_{cal} = 0.5223 < Z_{tab} = 1.96$ entonces no se rechaza H_0 , es decir, con un nivel de confiabilidad del 95% concluimos que el número de excedencias del ozono arriba del umbral 16 se distribuye de acuerdo a una distribución Poisson.

6.- CONCLUSIONES

- La estadística de prueba I bajo la hipótesis nula se distribuye asintóticamente normal con media 1 y varianza $\frac{2}{n}$, de donde se deduce que no depende del parámetro λ .

- La diferencia entre los cuantiles estimados de la estadística de prueba I mediante simulación por el Método Monte para los diferentes valores de λ en cada tamaño de muestra es mínima por lo que se reafirma que efectivamente la distribución de la estadística de prueba bajo la hipótesis nula no depende del parámetro λ . Además conforme aumentamos el tamaño de la muestra en ambos valores de λ la distribución de la estadística de prueba tiende a una distribución normal con media 1 y varianza $\frac{2}{n}$.

- Prácticamente los cuantiles estimados de la estadística de prueba I mediante simulación por el Método Monte Carlo son los mismos con los obtenidos directamente de la distribución $N\left(1, \frac{2}{n}\right)$.

-
-
- Estandarizando la estadística de prueba $I = \frac{\bar{X}}{S^2} \approx N\left(1, \frac{2}{n}\right)$, tenemos que $I^* = Z_{cal}$

$$= \sqrt{\frac{n}{2}} \left(\frac{\bar{X}}{S^2} - 1 \right) = T = t(x) \approx N(0,1)$$
, por lo que, para la prueba propuesta, la estadística de prueba es I^* .
 - Para valores de λ menores que 1, la probabilidad estimada del error tipo I cuando los datos pertenecen a la distribución Poisson es menor al nivel de significancia señalado. Por lo tanto como la probabilidad estimada del error tipo I es menor al nivel de significancia (α), en estos valores del parámetro de la distribución Poisson podemos decir que existe evidencia de que tenemos una prueba de tamaño α .
 - Para los diferentes parámetros considerados de las distribuciones Binomial y Binomial Negativa la potencia de la prueba estimada en promedio es mayor de 0.5 para todos los niveles de significancia y todos los tamaños de muestra. Además en todos los casos al aumentar el tamaño de la muestra, aumenta la potencia. Por lo tanto la prueba es consistente respecto al tamaño de la muestra
 - Los valores de la potencia de la prueba propuesta, para los diferentes tamaños de muestra como para los diferentes niveles de significancia, son mayores que los de la prueba de bondad de ajuste del índice de dispersión de Fisher cuando H_a : Binomial o Binomial Negativa con para diferentes parámetros, por lo que es mejor la prueba que se esta proponiendo.

7.- RECOMENDACIONES

Como se puede demostrar dicha prueba es muy simple de aplicar, para modelar fenómenos aleatorios en donde las variables aleatorias se refieren a conteos, por lo que se recomienda su difusión para su uso como una de las mejores pruebas de bondad de ajuste para la distribución Poisson, pues la potencia de esta prueba, tanto para los diferentes tamaños de muestra, como para los diferentes niveles de significancia son mayores que las de las pruebas de bondad de ajuste consideradas como comparación en esta investigación. Además con la finalidad de continuar ampliando su aplicación se recomienda investigar algunas extensiones de la misma.

8.- BIBLIOGRAFIA

Andrews, D. F. and Herzberg, A. M. (1985). Data: A collection of problems from many fields for the student and research worker. Springer-Verlag. New York, USA.

442 p.

Consul, P. C. and Jain, G. C. (1973). A generalization of the poisson distribution . Technometric, 15(4): 791-799.

Fisher, R.A. (1950). The significance of deviations from expectation in a Poisson series. Biometrics, 6: 17-24.

Hernández de la R.F.A. (1998). Una prueba de razón de verosimilitudes para Poisson contra Binomial Negativa. Tesis de Maestría. Colegio de Postgraduados, México. 73 p.

Lin, L. I-Kuei and Sanford, R. L. (1983). The robustness of the likelihood ratio test, the nonparametric sum rank test, and F-ratio test when the populations are from the negative binomial family . coomun. Statist.-Simulation and Computation, 12(5): 523-539.

Nakamura, M. and Pérez –Abreu,V. (1993). Exploratory data analysis for counts using the empirical probability generating function. Commun. Statist.-Theory and Methods, 22(3): 827-842.

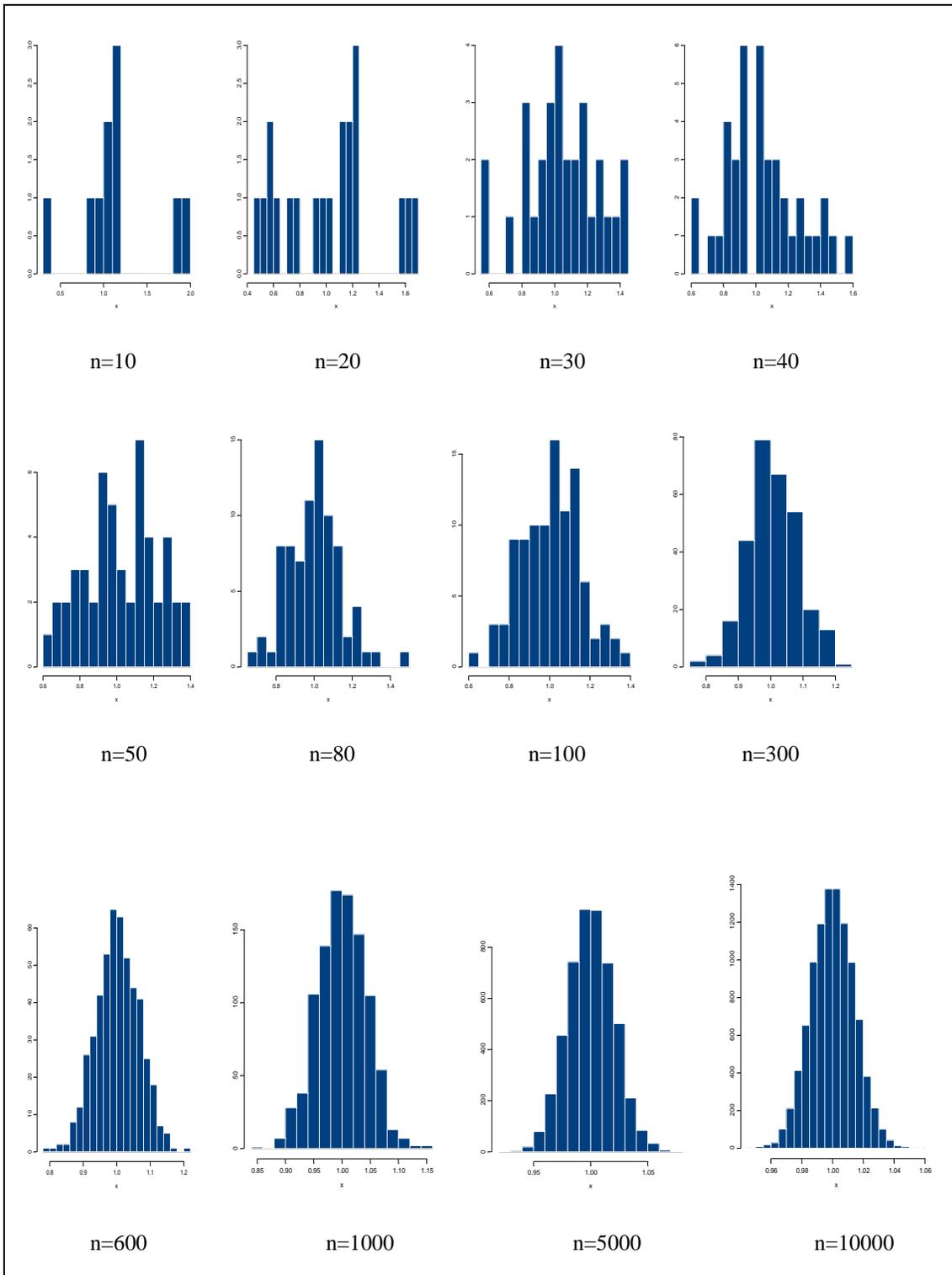
Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen in random sampling. *Philosophical Magazine. Series 5*, 50: 157-172.

Smith, R. L. (1989). Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone. *Statistical Science*, 4(4): 367-393.

Spinelli, J. J. and Stephens, M. A. (1997). Cramér-von Mises test of fit for the poisson distribution. *The Canadian Journal of Statistics*, 25(2): 257-268.

9.- ANEXOS

9.1.-ANEXO A

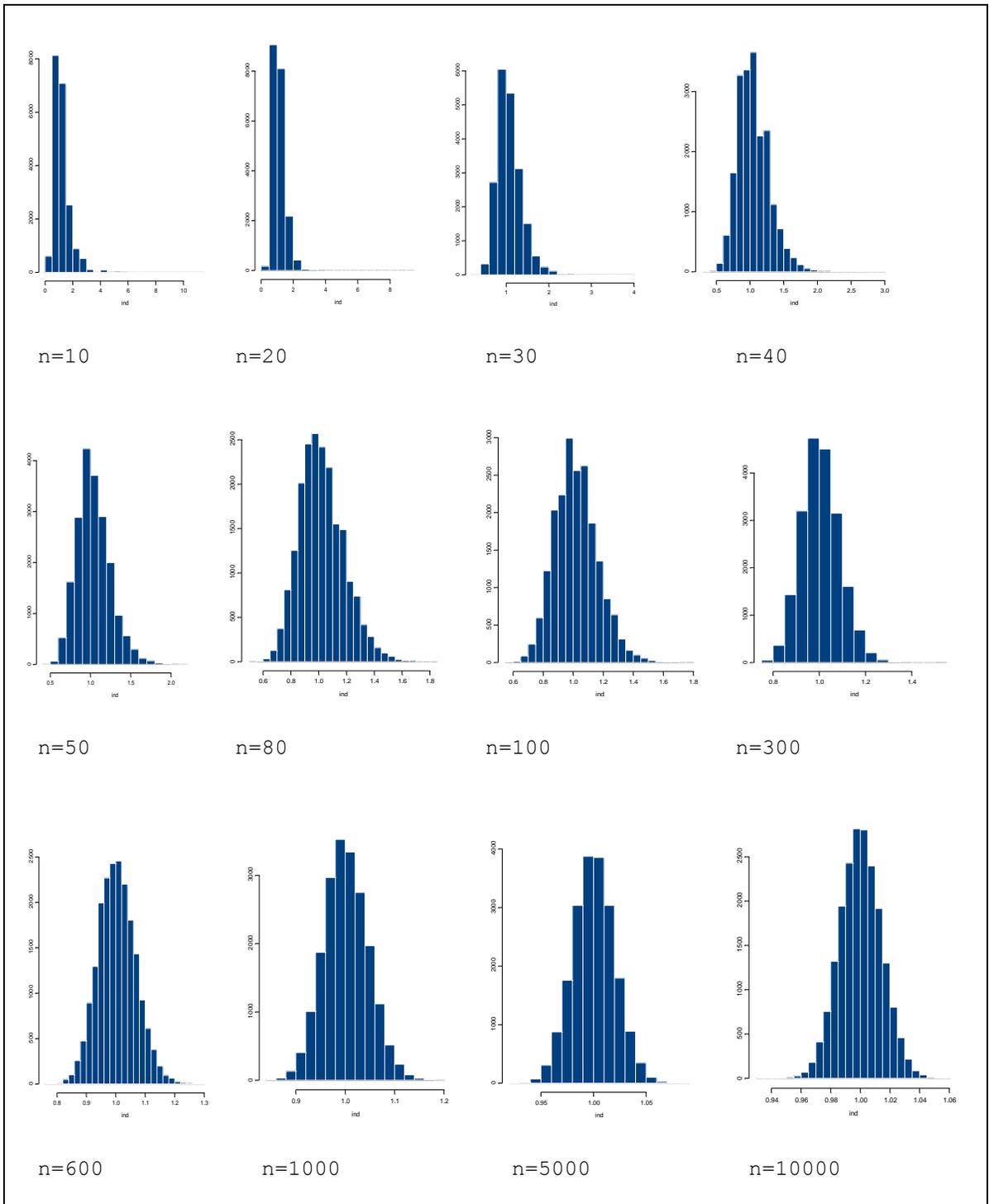


Gráfica 4.- Histogramas de la distribución $N\left(1, \frac{2}{n}\right)$

Tabla 8.- Cuantiles de la distribución $N\left(1, \frac{2}{n}\right)$.

n/α	0.0125	0.025	0.05	0.1	0.2	0.4	0.5
10	0.4518	0.5085	0.6220	0.8489	0.9768	1.0666	1.1018
20	0.4806	0.4957	0.5259	0.5573	0.6274	0.9558	1.0612
30	0.5699	0.5800	0.6477	0.8049	0.8754	1.0021	1.0477
40	0.6108	0.6209	0.7098	0.8016	0.8643	0.9404	1.0259
50	0.6531	0.6668	0.6937	0.7709	0.8457	0.9629	1.0233
60	0.6619	0.6908	0.7402	0.7793	0.8285	0.9739	1.0185
70	0.6900	0.7149	0.7504	0.8158	0.8842	0.9749	1.0154
80	0.7003	0.7226	0.8053	0.8184	0.8859	0.9726	1.0136
90	0.7190	0.7317	0.8032	0.8329	0.8960	0.9878	1.0125
100	0.7206	0.7491	0.7835	0.8132	0.8809	0.9826	1.0108
200	0.8142	0.8257	0.8570	0.8823	0.9334	0.9875	1.0054
300	0.8434	0.8541	0.8714	0.9082	0.9434	0.9854	1.0041
500	0.8673	0.8797	0.8959	0.9190	0.9493	0.9858	1.0022
1000	0.9039	0.9134	0.9311	0.9469	0.9631	0.9879	1.0004
5000	0.9562	0.9619	0.9672	0.9743	0.9831	0.9950	1.0002
10000	0.9690	0.9724	0.9768	0.9820	0.9879	0.9965	1.0000

n/α	0.5	0.6	0.8	0.9	0.95	0.975	0.9875
10	1.1018	1.1332	1.3122	1.8891	1.8946	1.8973	1.8987
20	1.0612	1.1411	1.2084	1.5651	1.6409	1.6501	1.6547
30	1.0477	1.0958	1.2235	1.3335	1.3847	1.4179	1.4288
40	1.0259	1.0602	1.2346	1.3850	1.4452	1.4852	1.5346
50	1.0233	1.1074	1.1857	1.2900	1.3126	1.3603	1.3745
60	1.0185	1.0490	1.1608	1.2373	1.2841	1.3710	1.4292
70	1.0154	1.0441	1.1113	1.2077	1.2802	1.3359	1.3753
80	1.0136	1.0394	1.1113	1.1652	1.2277	1.2657	1.3148
90	1.0125	1.0420	1.1137	1.1708	1.1963	1.3027	1.3140
100	1.0108	1.0476	1.1298	1.1897	1.2547	1.3051	1.3343
200	1.0054	1.0318	1.0893	1.1219	1.1502	1.1879	1.2271
300	1.0041	1.0289	1.0695	1.1066	1.1450	1.1649	1.1781
500	1.0022	1.0182	1.0582	1.0868	1.1080	1.1249	1.1431
1000	1.0004	1.0122	1.0372	1.0538	1.0689	1.0788	1.0965
5000	1.0002	1.0052	1.0175	1.0262	1.0331	1.0398	1.0459
10000	1.0000	1.0037	1.0119	1.0181	1.0232	1.0273	1.0313

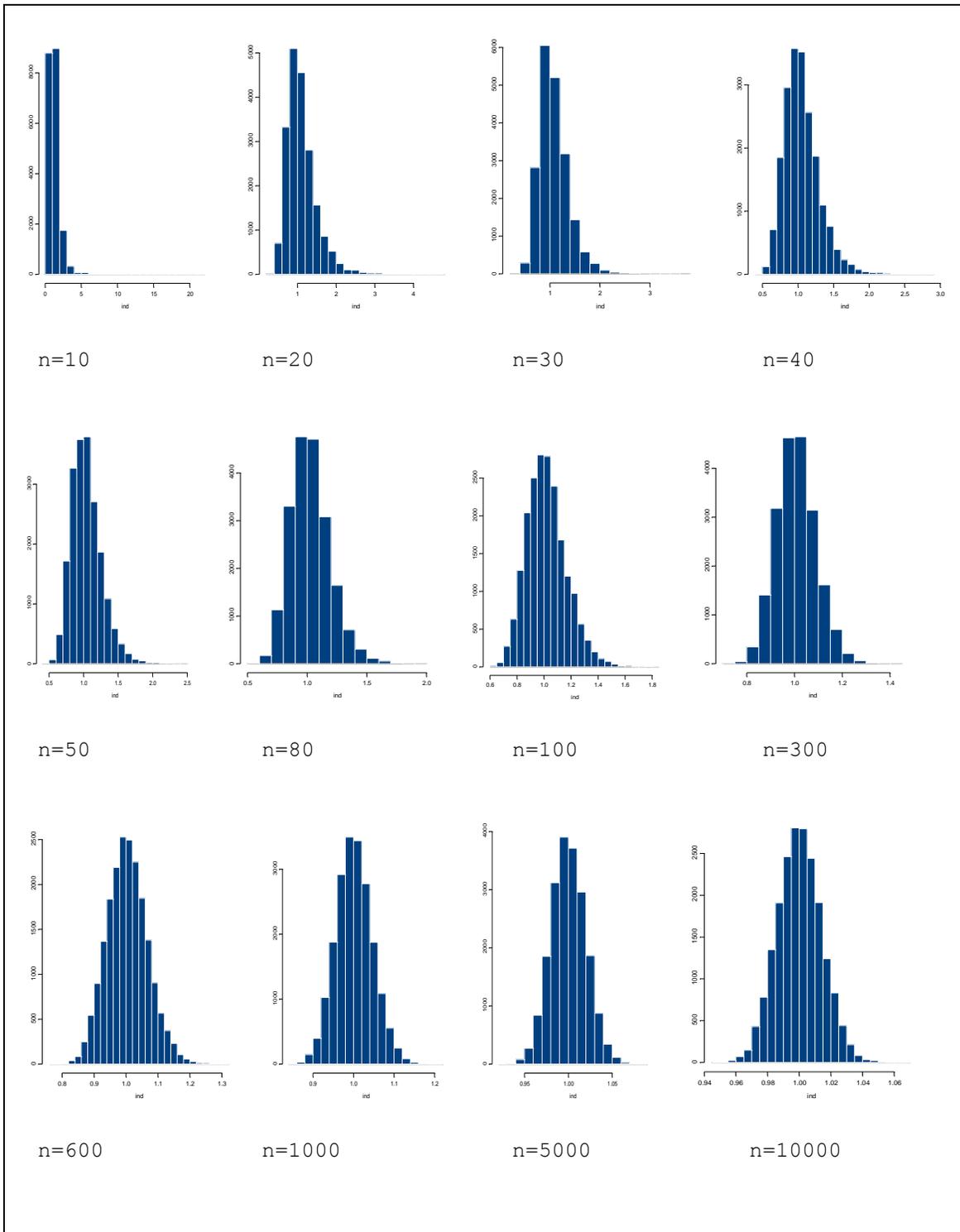


Gráfica 5.- Histogramas de la estadística de prueba I bajo H_0 de x se distribuye $Poisson(\lambda = 1)$, estimados con 20000 muestras, por el método Monte Carlo.

Tabla 9.- Cuantiles de la distribución de la estadística de prueba I bajo H_0 de que x se distribuye Poisson($\lambda=1$), estimados con 20000 muestras, por el método Monte Carlo

n / α	0.0125	0.025	0.05	0.1	0.2	0.4	0.5
10	0.4323	0.4792	0.5436	0.6428	0.7500	0.9473	1.1123
20	0.5277	0.5738	0.6333	0.7161	0.7995	0.9587	1.0555
30	0.5838	0.6304	0.6842	0.7512	0.8423	0.9707	1.0357
40	0.6242	0.6678	0.7138	0.7734	0.8502	0.9756	1.0282
50	0.6546	0.6893	0.7332	0.7874	0.8688	0.9778	1.0208
60	0.6757	0.7084	0.7488	0.8062	0.8695	0.9700	1.0172
70	0.6969	0.7337	0.7727	0.8179	0.8828	0.9668	1.0153
80	0.7181	0.7473	0.7814	0.8288	0.8854	0.9682	1.0128
90	0.7218	0.7529	0.7887	0.8347	0.8908	0.9704	1.0113
100	0.7362	0.7650	0.8033	0.8466	0.8982	0.9737	1.0110
200	0.8048	0.8284	0.8540	0.8861	0.9254	0.9798	1.0052
300	0.8379	0.8562	0.8784	0.9039	0.9358	0.9835	1.0034
500	0.8686	0.8842	0.9029	0.9235	0.9494	0.9861	1.0021
1000	0.9061	0.9173	0.9308	0.9455	0.9641	0.9891	1.0006
5000	0.9564	0.9617	0.9675	0.9748	0.9836	0.9951	1.0002
10000	0.9690	0.9729	0.9774	0.9822	0.9883	0.9967	1.0001

n/α	0.5	0.6	0.8	0.9	0.95	0.975	0.9875
10	1.1123	1.1739	1.5365	2.0000	2.2500	2.8643	3.2727
20	1.0555	1.1099	1.3571	1.5405	1.7448	1.9000	2.1326
30	1.0357	1.1112	1.2910	1.4292	1.5675	1.7058	1.8730
40	1.0282	1.0833	1.2284	1.3623	1.4819	1.5839	1.6922
50	1.0208	1.0675	1.2043	1.3095	1.4044	1.5076	1.5981
60	1.0172	1.0586	1.1800	1.2772	1.3691	1.4538	1.5526
70	1.0153	1.0537	1.1694	1.2602	1.3440	1.4184	1.4874
80	1.0128	1.0477	1.1554	1.2343	1.3071	1.3760	1.4351
90	1.0113	1.0461	1.1410	1.2231	1.2918	1.3576	1.4127
100	1.0110	1.0515	1.1394	1.2073	1.2730	1.3293	1.3821
200	1.0052	1.0295	1.0924	1.1408	1.1834	1.2227	1.2594
300	1.0034	1.0240	1.0740	1.1133	1.1475	1.1787	1.2040
500	1.0021	1.0184	1.0566	1.0874	1.1121	1.1330	1.1534
1000	1.0006	1.0120	1.0393	1.0598	1.0765	1.0925	1.1065
5000	1.0002	1.0052	1.0167	1.0255	1.0331	1.0397	1.0453
10000	1.0001	1.0037	1.0120	1.0183	1.0237	1.0281	1.0320

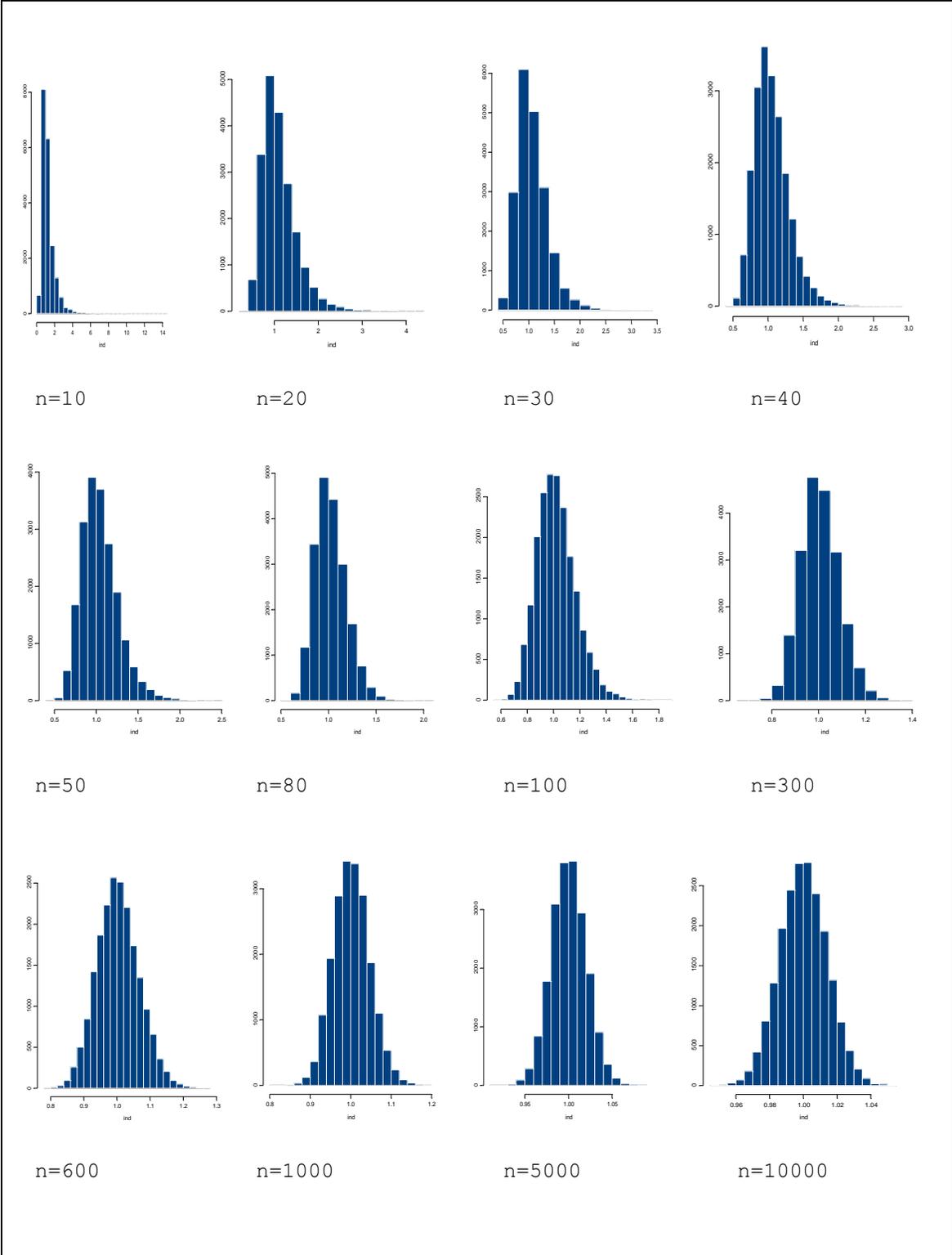


Gráfica 6.- Histogramas de la estadística de prueba I bajo H_0 de que x se distribuye Poisson($\lambda = 2$), estimados con 20000 muestras, por el método Monte Carlo

Tabla 10.- Cuantiles de la distribución de la estadística de prueba I bajo H_0 de que x se distribuye Poisson($\lambda=2$), estimados con 20000 muestras, por el método Monte Carlo

n / α	0.0125	0.025	0.05	0.1	0.2	0.4	0.5
10	0.4285	0.4759	0.5415	0.6265	0.7500	0.9503	1.0800
20	0.5224	0.5716	0.6272	0.6989	0.7964	0.9545	1.0301
30	0.5918	0.6355	0.6826	0.7461	0.8307	0.9666	1.0275
40	0.6290	0.6673	0.7142	0.7727	0.8478	0.9647	1.0211
50	0.6567	0.6922	0.7354	0.7877	0.8592	0.9631	1.0169
60	0.6870	0.7218	0.7596	0.8071	0.8715	0.9710	1.0176
70	0.7000	0.7347	0.7707	0.8190	0.8816	0.9720	1.0150
80	0.7172	0.7469	0.7848	0.8295	0.8871	0.9717	1.0128
90	0.7270	0.7576	0.7946	0.8369	0.8911	0.9687	1.0053
100	0.7407	0.7675	0.8008	0.8411	0.8937	0.9714	1.0058
200	0.8018	0.8261	0.8527	0.8841	0.9229	0.9786	1.0041
300	0.8380	0.8569	0.8781	0.9050	0.9376	0.9835	1.0034
500	0.8718	0.8855	0.9024	0.9233	0.9497	0.9861	1.0019
1000	0.9045	0.9176	0.9303	0.9456	0.9640	0.9893	1.0003
5000	0.9568	0.9622	0.9682	0.9752	0.9832	0.9947	0.9998
10000	0.9700	0.9732	0.9774	0.9822	0.9883	0.9965	1.0000

n/α	0.5	0.6	0.8	0.9	0.95	0.975	0.9875
10	1.0800	1.1911	1.5882	2.0769	2.5596	3.0000	3.7317
20	1.0301	1.1176	1.3571	1.5833	1.8115	2.0283	2.2848
30	1.0275	1.0943	1.2852	1.4448	1.6175	1.7625	1.9409
40	1.0211	1.0822	1.2319	1.3741	1.5000	1.6397	1.7588
50	1.0169	1.0680	1.2111	1.3243	1.4375	1.5415	1.6449
60	1.0176	1.0647	1.1848	1.2902	1.3840	1.4642	1.5470
70	1.0150	1.0577	1.1674	1.2609	1.3419	1.4193	1.4992
80	1.0128	1.0529	1.1581	1.2443	1.3195	1.3974	1.4695
90	1.0053	1.0451	1.1413	1.2227	1.2978	1.3609	1.4235
100	1.0058	1.0421	1.1356	1.2157	1.2780	1.3408	1.3987
200	1.0041	1.0288	1.0925	1.1419	1.1870	1.2239	1.2651
300	1.0034	1.0243	1.0749	1.1150	1.1492	1.1779	1.2061
500	1.0019	1.0182	1.0563	1.0858	1.1122	1.1360	1.1552
1000	1.0003	1.0117	1.0390	1.0600	1.0778	1.0931	1.1073
5000	0.9998	1.0050	1.0169	1.0258	1.0329	1.0397	1.0461
10000	1.0000	1.0035	1.0117	1.0183	1.0235	1.0280	1.0321



Gráfica 7.- Histogramas de la estadística de prueba I bajo H_0 de que x se distribuye Poisson($\lambda = 3$), estimados con 20000 muestras, por el método Monte Carlo

Tabla 11.- Cuantiles de la distribución de la estadística de prueba I bajo H_0 de que x se distribuye Poisson($\lambda=3$), estimados con 20000 muestras, por el método Monte Carlo

n / α	0.0125	0.025	0.05	0.1	0.2	0.4	0.5
10	0.4285	0.4736	0.5400	0.6174	0.7467	0.9642	1.0673
20	0.5347	0.5740	0.6309	0.7030	0.7970	0.9564	1.0384
30	0.5897	0.6297	0.6780	0.7395	0.8244	0.9586	1.0204
40	0.6330	0.6689	0.7133	0.7697	0.8470	0.9611	1.0183
50	0.6590	0.6929	0.7341	0.7893	0.8617	0.9688	1.0187
60	0.6823	0.7159	0.7574	0.8092	0.8734	0.9681	1.0127
70	0.7038	0.7359	0.7728	0.8194	0.8800	0.9692	1.0119
80	0.7153	0.7473	0.7837	0.8260	0.8830	0.9666	1.0068
90	0.7319	0.7598	0.7927	0.8377	0.8916	0.9716	1.0096
100	0.7411	0.7695	0.8000	0.8441	0.8962	0.9717	1.0091
200	0.8075	0.8287	0.8552	0.8865	0.9248	0.9790	1.0040
300	0.8396	0.8568	0.8786	0.9046	0.9369	0.9822	1.0029
500	0.8702	0.8849	0.9030	0.9238	0.9496	0.9859	1.0014
1000	0.9065	0.9192	0.9304	0.9451	0.9636	0.9892	1.0010
5000	0.9562	0.9621	0.9680	0.9752	0.9835	0.9950	1.0002
10000	0.9692	0.9731	0.9771	0.9821	0.9880	0.9964	1.0000

n/α	0.5	0.6	0.8	0.9	0.95	0.975	0.9875
10	1.0673	1.2183	1.6508	2.1235	2.6052	3.2586	3.8571
20	1.0384	1.1301	1.3775	1.6219	1.8387	2.0821	2.3298
30	1.0204	1.0935	1.2847	1.4474	1.6111	1.7731	1.9519
40	1.0183	1.0791	1.2392	1.3777	1.5075	1.6374	1.7614
50	1.0187	1.0718	1.2095	1.3290	1.4491	1.5529	1.6630
60	1.0127	1.0608	1.1831	1.2849	1.3895	1.4807	1.5637
70	1.0119	1.0561	1.1694	1.2641	1.3502	1.4315	1.5068
80	1.0068	1.0497	1.1567	1.2439	1.3220	1.3881	1.4546
90	1.0096	1.0468	1.1469	1.2270	1.3001	1.3674	1.4355
100	1.0091	1.0456	1.1383	1.2137	1.2841	1.3410	1.4020
200	1.0040	1.0293	1.0926	1.1436	1.1872	1.2274	1.2642
300	1.0029	1.0242	1.0757	1.1146	1.1487	1.1776	1.2046
500	1.0014	1.0170	1.0568	1.0872	1.1109	1.1333	1.1525
1000	1.0010	1.0124	1.0390	1.0600	1.0775	1.0925	1.1067
5000	1.0002	1.0053	1.0173	1.0263	1.0338	1.0405	1.0465
10000	1.0000	1.0034	1.0119	1.0180	1.0233	1.0275	1.0313

9.2.-ANEXO B

Programa 1.- Para la obtención de los cuantiles de la distribución $N\left(1, \frac{2}{n}\right)$

```
x<-rnorm(100,1,0.02)
hist(x,nclass=20)
q1<-quantile(x,c(0.0125,0.025,0.05,0.1,,0.2,0.4,0.5,))
q2<-quantile(x,c(0.5,0.6,0.8,0.9,0.95,0.975,0.9875))
q1
q2
```

Programa 2.- Para la obtención de los cuantiles de la distribución de la estadística de prueba I bajo H_0 de que x se distribuye Poisson($\lambda=1$), estimados con 20000 muestras, por el método Monte Carlo

```
for(i in 1:20000)
{
  ind<-matrix(nrow=20000,ncol=1)
  x<-rpois(10,1)
  y[i]<-mean(x)/var(x)
  ind<-y
}
hist(ind,nclass=20)
q1<-quantile(ind,c(0.0125,0.025,0.05,0.1,,0.2,0.4,0.5,))
q2<-quantile(ind,c(0.5,0.6,0.8,0.9,0.95,0.975,0.9875))
q1
q2
```

Programa 3.- Para la obtención del tamaño de la prueba para la distribución

Poisson($\lambda = 0.2$)

```
for(i in 1:10000)
{
  ind<-matrix(nrow=10000,ncol=1)
  x<-rpois(30,0.2)
  y[i]<-(sqrt(30/2))*((mean(x)/var(x))-1)
  ind<-y
}
cont<-as.matrix(ind[ind>-1.96])
cont1<-as.matrix(cont[cont>=1.96])
tam<-10000-nrow(cont)
tam1<-nrow(cont1)
tp<-(tam+tam1)/10000
tp
```

Programa 4.- Para la obtención de la potencia de la prueba cuando la H_a es una

Binomial con media 1 usando los valores de la distribución normal estándar

```
for(i in 1:10000)
{
  ind<-matrix(nrow=10000,ncol=1)
  x<-rbinom(30,2,0.5) # sample of size 10 with mean 2*0.5 = 1
  y[i]<-(sqrt(30/2))*((mean(x)/var(x))-1)
  ind<-y
}
```

```

cont<-as.matrix(ind[ind>-1.96])
cont1<-as.matrix(cont[cont>=1.96])
tam<-10000-nrow(cont)
tam1<-nrow(cont1)
p<-(tam+tam1)/10000
p

```

Programa 5.- Para la obtención de la potencia de la prueba cuando la H_a es una Binomial con media 1 usando los valores de la distribución ji-cuadrada

```

for(i in 1:10000)
{
  r<-matrix(nrow=10000,ncol=1)
  x<-rbinom(20,2,0.5) # sample of size 20 with mean 2*0.5 = 1
  y[i]<-(20*(var(x)))/(mean(x))
  r<-y
}
cont<-as.matrix(r[r<8.91])
tam<-nrow(cont)
p<-tam/10000
p

```