



Súper nota

**Nombre de la alumna: Jazmín Escobedo
Gómez**

Parcial: Cuarto

Materia: Bioestadística

**Nombre del maestro: Lic. Icel Bernardo
Lepe Arriaga**

Cuatrimestre: Cuarto

Licenciatura: Enfermería

**Frontera comlapa Chiapas a 28 de
noviembre del 2024**

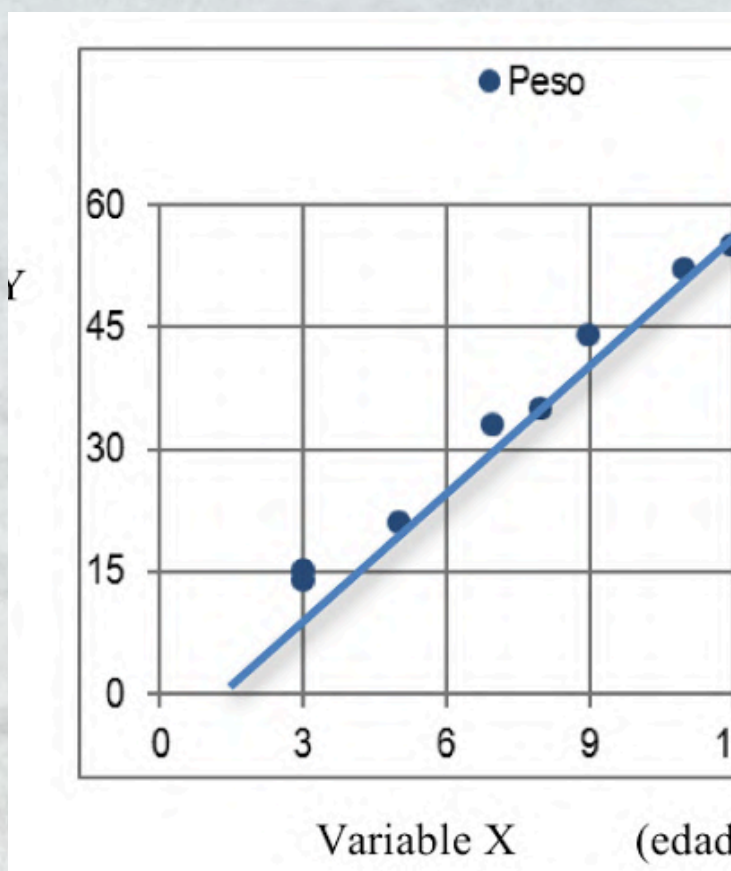
Test para poblaciones normales

Una población

En la inferencia sobre una variable numérica en una población, el objetivo principal de los test de hipótesis es contrastar el valor de alguna medida de posición (media o mediana), de dispersión (varianza) o de algún otro parámetro poblacional.



En principio, si hacemos inferencia sobre un resumen de la variable podemos considerar que dicha variable sigue una distribución y plantear el problema desde la óptica paramétrica. Esa distribución puede ser cualquiera de las muchas distribuciones tipo (binomial, Poisson, normal, uniforme, gamma,)



En la inferencia sobre dos variables numéricas en una población se trata de ver si existe relación lineal entre las mismas a partir de la información muestral. Así, podemos tratar de analizar si existe relación entre la renta y el consumo de las familias y cuál es la intensidad y el sentido de la misma,

DISTRIBUCION GAMMA

Propiedades de la función gamma

Esta función también se la puede representar como:

mostracion: $\Gamma(\alpha) = (\alpha - 1)!$

integrando por partes:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$$

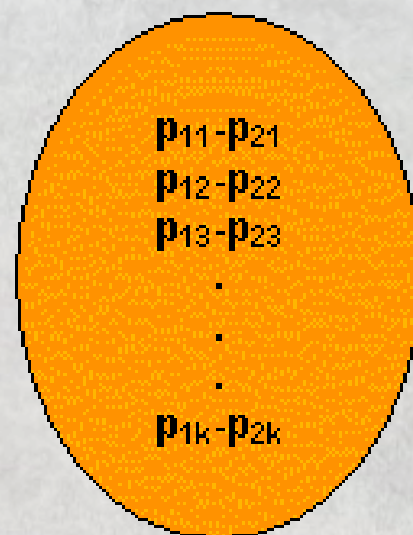
$$u = x^{\alpha-1} \quad dv = e^{-x} dx$$

$$du = (\alpha-1)x^{\alpha-2} dx \quad v = -e^{-x}$$

Si podemos aceptar normalidad, el objetivo general de comparar dos poblaciones se traduce en comparar las medias de la variable en cada una de ellas; aunque suele tener únicamente un interés instrumental, también se pueden comparar las varianzas.

Dos poblaciones

En este caso, el objetivo fundamental es comparar la distribución de una variable cuantitativa X en las dos poblaciones (subpoblaciones) determinadas por las modalidades de una característica cualitativa dicotómica o, lo que es lo mismo, estudiar si la variable cuantitativa



Distribución muestral de Diferencia de Proporciones

Para poder realizar estas comparaciones se utilizan dos muestras provenientes de individuos diferentes (estudiar las posibles diferencias salariales en función del sexo a partir de una muestra de hombres y una muestra de mujeres); en este caso hablaremos de muestras independientes

Comparación de medias en poblaciones normales.

Si $\sigma_1^2 = \sigma_2^2 = \sigma^2 \Rightarrow S^2 = \frac{S_1^2 + S_2^2}{2} \Rightarrow S_d = \sqrt{\frac{S^2}{N_1} + \frac{S^2}{N_2}}$

Si H_0 es cierta $\rightarrow m_1 - m_2 = 0$

Se tiene que:

$$\frac{d - E(d)}{S_d} = \frac{\bar{X}_1 - \bar{X}_2 - (m_1 - m_2)}{\sqrt{\frac{S^2}{N_1} + \frac{S^2}{N_2}}} \approx t_{(N_1-1)+(N_2-1)}$$

$$\frac{\bar{X}_1 - \bar{X}_2}{S_d} \approx t_{(N_1-1)+(N_2-1)}$$

Suponga una distribución aproximadamente normal.

SOLUCIÓN:

$$\mu = \bar{x} \pm \frac{t_{\alpha/2}}{2} (S/\sqrt{n})$$

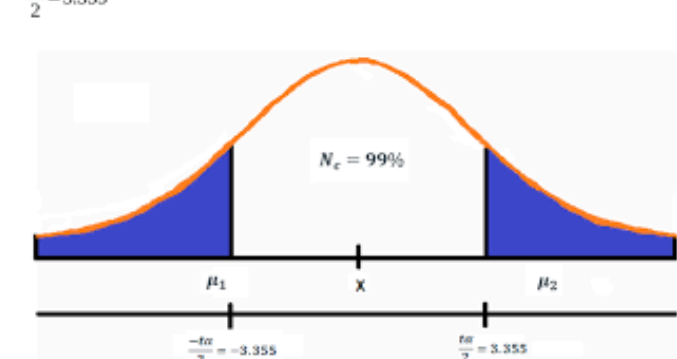
$$n=9$$

$$V=8$$

$$\bar{x}=1.0055$$

$$S=0.02455$$

$$\frac{t_{\alpha/2}}{2}=3.355$$



$$\mu = \begin{cases} 0.978 \\ 1.032 \end{cases}$$

$$e=0.027 \text{ cm}$$

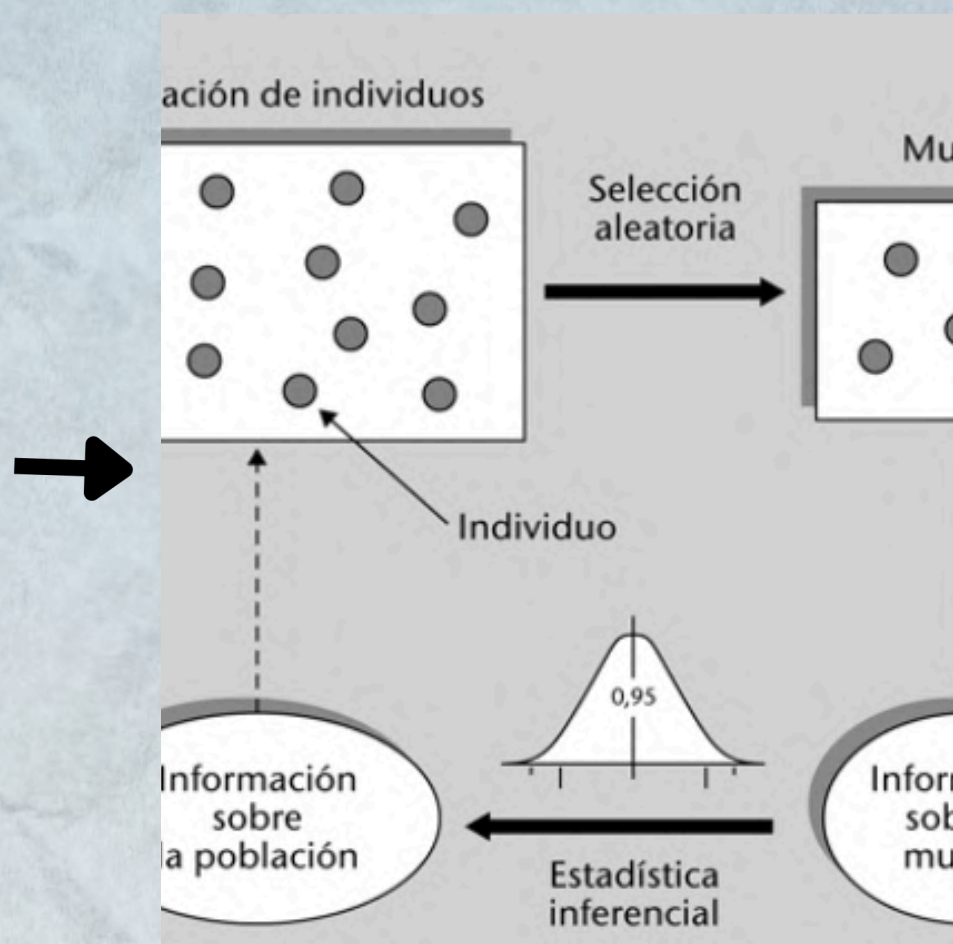
La máquina está produciendo cilindros con un diámetro entre 0.978cm y 1.032cm con un nivel de confianza del 99% y con un error de 0.027cm.

Test para poblaciones binomiales y de Poisson

Nos encontramos con un modelo derivado de un proceso experimental puro, en el que se plantean las siguientes circunstancias.

Se realiza un número n de pruebas (separadas o separables).

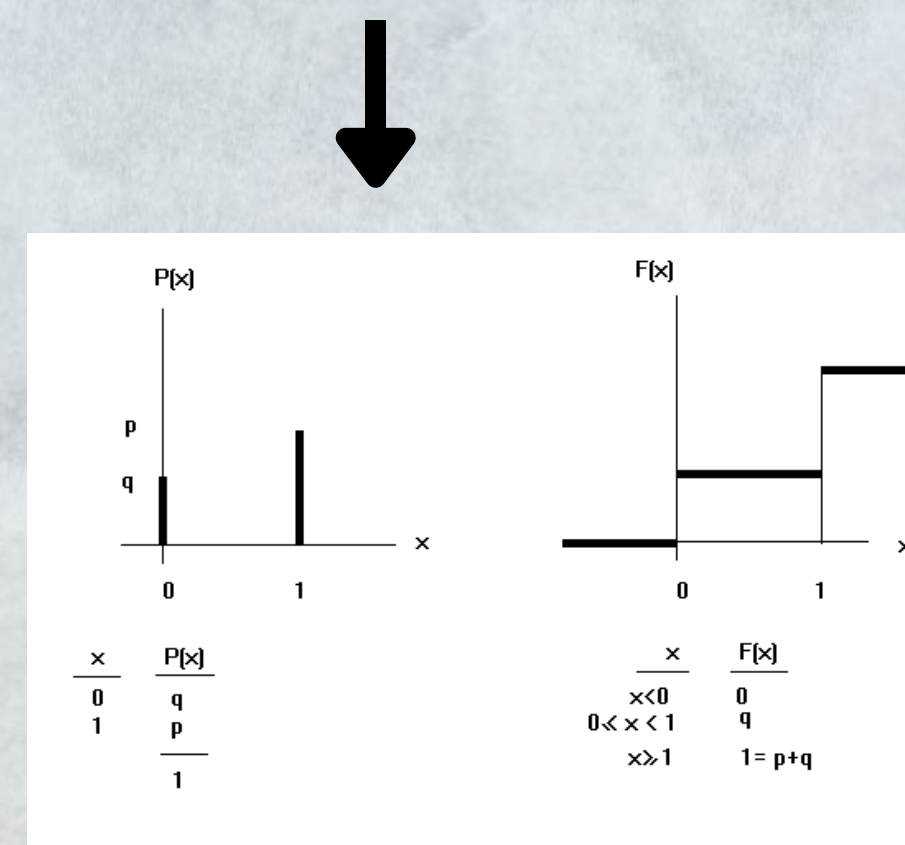
Cada prueba puede dar dos únicos resultados A y \bar{A} .



La probabilidad de obtener un resultado A es p y la de obtener un resultado \bar{A} es q , con $q = 1 - p$, en todas las pruebas. Esto implica que las pruebas se realizan exactamente en las mismas condiciones y son, por tanto, independientes en sus resultados. Si se trata de extracciones, (muestreo), las extracciones deberán ser con devolución

- 2 Número fijo de ensayos
- 3 La misma probabilidad de éxito
- 4 Gran tamaño de muestra

Un uso común del test binomial es en el caso donde la hipótesis nula es aquella en la que las dos categorías son igualmente probables de que ocurran (como el lanzamiento de una moneda). Las tablas están ampliamente disponibles para dar la importancia observada en el número



Para muestras grandes como las del siguiente ejemplo, La distribución binomial es bien aproximada por convenientes distribuciones continuas, y éstos se utilizan como la base para las pruebas alternativas que son mucho más rápidas para computar, Prueba χ^2 de Pearson y el G-test.

agrupados

datos de 21 estudiantes están dados por la siguiente

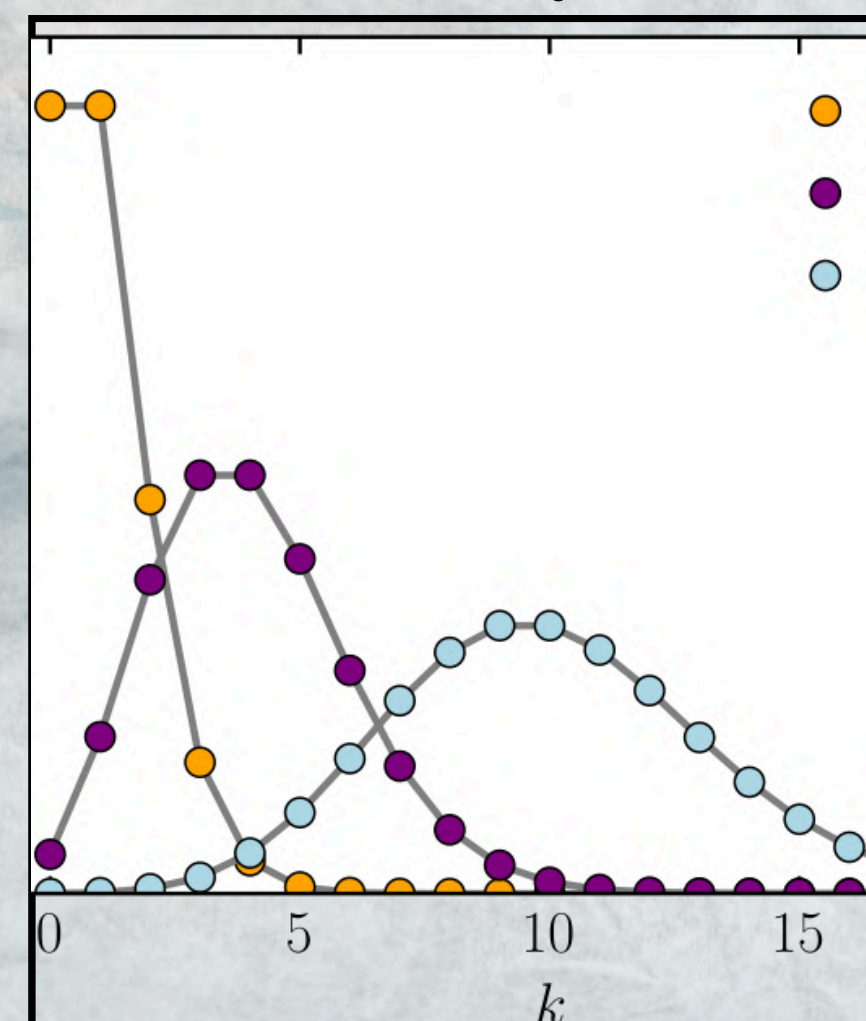
Interval	x_i	f_i
[55,59)	57	2
[59,63)	61	5
[63,67)	65	3
[67,71)	69	7
[71,75)	73	4

$\bar{x} = \frac{57.2 + 61.5 + 65.8}{3} = 66.142$

La distribución de Poisson se aplica a varios fenómenos discretos de la naturaleza (esto es, aquellos fenómenos que ocurren 0, 1, 2, 3... veces durante un periodo definido de tiempo o en un área determinada) cuando la probabilidad de ocurrencia del fenómeno es constante en el tiempo o el espacio.

El número de autos que pasan a través de un cierto punto en una ruta (suficientemente distantes de los semáforos) durante un periodo definido de tiempo.

- El número de errores de ortografía que uno comete al escribir una única página.
- El número de llamadas telefónicas en una central telefónica por minuto.



Probabilidad y Estadística

Variable Aleatoria Discreta

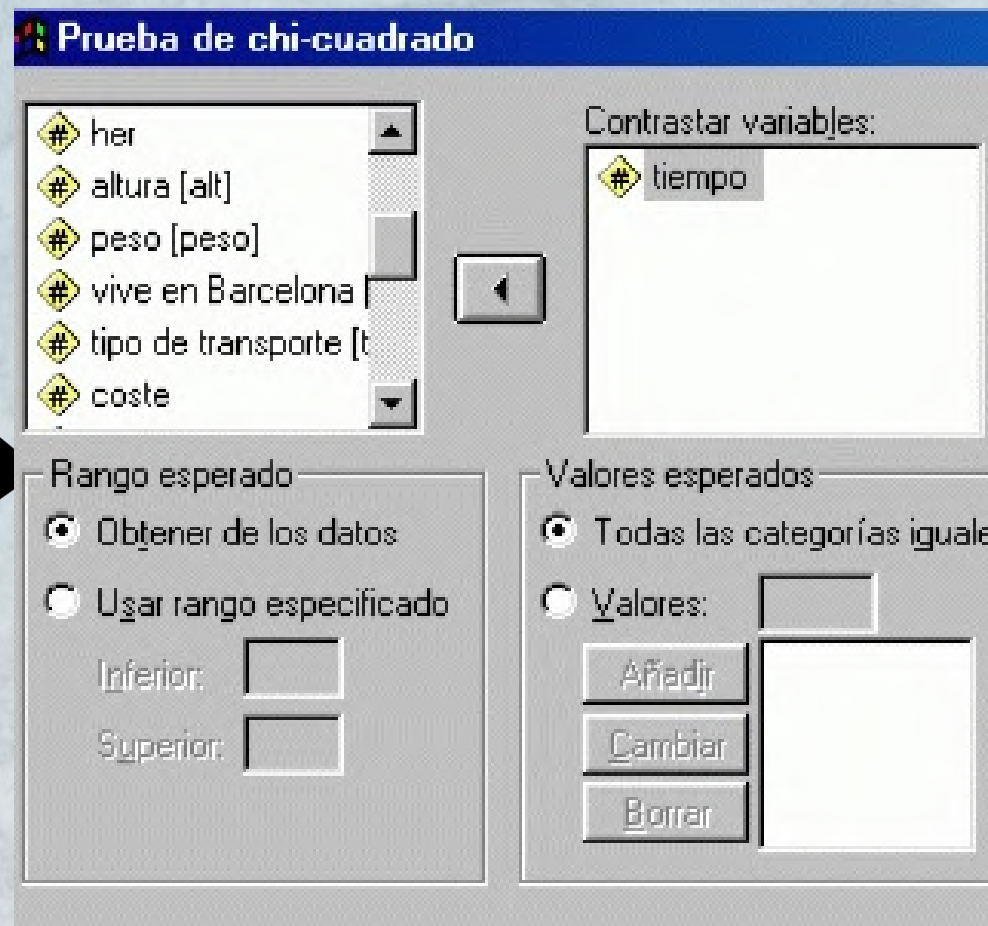
Distribución de Poisson

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

Profesor: CARLOS TEJADA

Test basado en el estadístico Chi-cuadrado

Esta prueba puede utilizarse incluso con datos medibles en una escala nominal. La hipótesis nula de la prueba Chi-cuadrado postula una distribución de probabilidad totalmente especificada como el modelo matemático de la población que ha generado la muestra.



La hipótesis ortostática Para realizar este contraste se disponen los datos en una tabla de frecuencias. Para cada valor o intervalo de valores se indica la frecuencia absoluta observada o empírica (O_i). A continuación, y suponiendo que la hipótesis nula es cierta

$$\sum O_i$$

Este estadístico tiene una distribución Chi-cuadrado con $k-1$ grados de libertad si n es suficientemente grande, es decir, si todas las frecuencias esperadas son mayores que 5. En la práctica se tolera un máximo del 20% de frecuencias inferiores a 5.

$$\sum \frac{(f_o - f_e)^2}{f_e}$$

frecuencia del valor observado
frecuencia del valor esperado

En Valores esperados se debe especificar la distribución teórica activando una de las dos alternativas. Por defecto está activada Todas las categorías iguales que recoge la hipótesis de que la distribución de la población es uniforme discreto

Si existe concordancia perfecta entre las frecuencias observadas y las esperadas el estadístico tomará un valor igual a 0; por el contrario, si existe una gran discrepancia entre estas frecuencias el estadístico tomará un valor grande y, en consecuencia, se rechazará la hipótesis nula.

$$\sum_{k=1}^n \frac{(O_k - E_k)^2}{E_k}$$

El cuadro de diálogo al que se accede con el botón Opciones ofrece la posibilidad de calcular los Estadísticos Descriptivos y/o los Cuartiles, así como seleccionar la forma en que se desea tratar los valores perdidos.

$$= \sum \frac{(O - E)^2}{E}$$

O = the test statistic
O = Observed frequencies E = Expected frequencies

$$\sum \frac{e_{(N_i+1)}}{e_{(N_i+1)}}$$

ésimo residual
tero más grande

Test de bondad de ajuste.

La bondad de ajuste de un modelo estadístico describe lo bien que se ajusta un conjunto de observaciones. Las medidas de bondad en general resumen la discrepancia entre los valores observados y los valores esperados en el modelo de estudio.

PRUEBAS DE BONDAD DE AJUSTE

Lo bien que se puede ajustar los datos u observaciones, y saber si los valores en o asemejan a lo esperado.

Estas pruebas tienen las siguientes pruebas de hipótesis:
 Ho: La muestra proviene de la distribución de interés.
 H1: La muestra no proviene de la distribución de interés.

De acuerdo a ello, se puede medir o cuantificar la semejanza entre las dos distribuciones, y ello se puede realizar a través de un estadístico que sigue una distribución Chi-Cuadrado, ya que mide la discrepancia o diferencia entre lo observado y los esperados.

En estas pruebas, se utiliza el valor crítico de un Chi-Cuadrado, donde:
 Se acepta Ho si: $X_0^2 > X_{\alpha, k-t-1}^2$ No se rechaza Ho si: $X_0^2 < X_{\alpha, k-t-1}^2$

INSTITUTO CIENTÍFICO DEL PACÍFICO

Estas pruebas permiten verificar que la población de la cual proviene una muestra tiene una distribución especificada o supuesta.

Sea X: variable aleatoria poblacional $f_0(x)$ la distribución (o densidad) de probabilidad especificada o supuesta para X.

Se desea probar la hipótesis: $H_0: f(x) = f_0(x)$



Se dice que la variable aleatoria continua X tiene distribución chi-cuadrado con r grados de libertad, y se denota por $X \sim \chi^2(r)$, si su función de densidad es

$$f(x) = \begin{cases} \frac{1}{\Gamma(\frac{r}{2})} x^{\frac{r}{2}-1} e^{-x/2}, & \text{si } x \geq 0 \\ 0, & \text{si } x < 0 \end{cases}$$

Donde r es un número entero positivo. Su gráfica, para distintos valores de los grados de libertad se muestra

NOTA:
 Si $X \sim \chi^2(r)$, esto es, si $X \sim r(1/2, 1/2)$, entonces:
 a) Su media es $\mu=r$
 b) Su varianza es $\sigma^2=2r$

PROPIEDADES DE LA DISTRIBUCIÓN CHI-CUADRADO
 Las siguientes propiedades (cuya prueba omitimos) son de importancia para el desarrollo de algunas aplicaciones.

Prueba de bondad de ajuste chi cuadrado χ^2

El procedimiento de la prueba requiere una muestra aleatoria de tamaño n proveniente de la población cuya distribución de probabilidad es desconocida.

Si la variable es continua se utiliza de modelo matemático de la distribución propuesta y se evalúan todas las categorías y luego se multiplica por el número total de datos. 4. Se calcula el estadístico de prueba



Observado

número de parámetros estimados

número de la muestra

número de categorías o clases

probabilidad

$$\sum \frac{(O_i - E_i)^2}{E_i}$$

La prueba se puede sintetizar en los siguientes pasos.

1. Se colocan los n datos históricos (muestrales) en una tabla de frecuencia.
2. Se propone una distribución de probabilidad una distribución de probabilidad de acuerdo con la tabla de frecuencia o con la curva que muestre un histograma o polígono de frecuencia.

PRUEBAS DE BONDAD DE AJUSTE Y PRUEBAS NO PARAMÉTRICAS

zona de aceptación $\chi^2_{\alpha, m-k-1}$
 zona de rechazo α

el estimador C es menor o igual al valor correspondiente X^2 con $m-k-1$ grados de libertad ($K =$ números de parámetros estimados de la distribución propuesta estimada por los estadísticos muestrales) y a un nivel de confiabilidad de $1-\alpha$, entonces no se puede rechazar la hipótesis de que los datos siguen la distribución que se propuso.

Prueba para la Bondad de ajuste

Región de aceptación $\alpha=0.05$
 Región de rechazo $\alpha=0.05$
 $\chi^2_{0.05, 9} = 16.92$

$\frac{(44-30)^2}{30}$	$\frac{(40-30)^2}{30}$	$\frac{(20-30)^2}{30}$	$\frac{(44-30)^2}{30}$	$\frac{(25-30)^2}{30}$
16.9	3.62	3.67	16.9	2.5

