UNIVERSIDAD NACIONAL DEL CHIMBORAZO

ANÁLISIS ESTOCÁSTICO PARTE III

Patricia Hernández, PhD

Asignatura

Modelización Económica

QUÉ HACER CON LOS DATOS DE UNA MUESTRA

Estimación

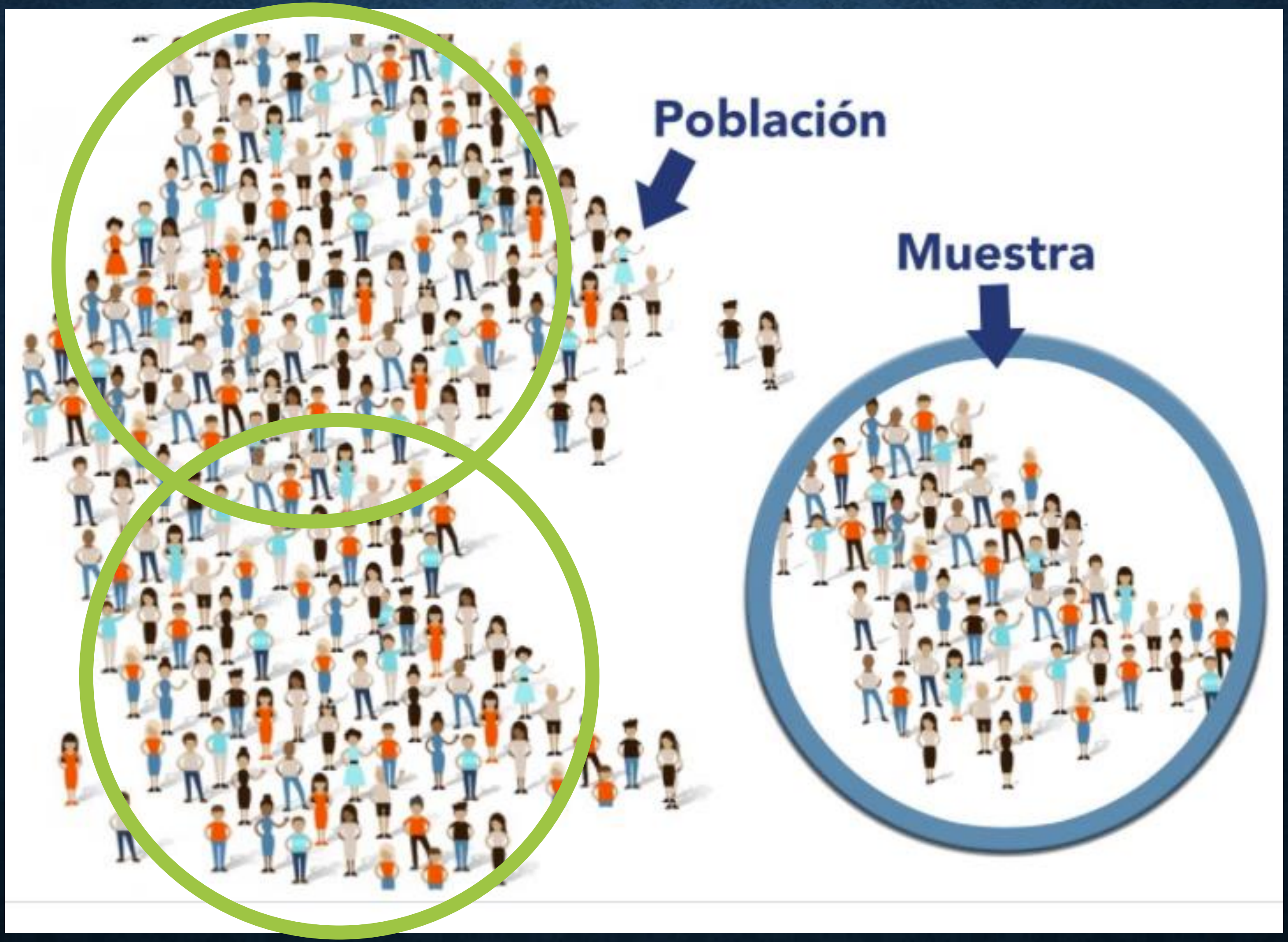
pone énfasis en
intentar cuantificar el tamaño del efecto detectado
(valor del parámetro)
(cuando difieren dos grupos, como de intensa es
a relación entre dos variables.

Prueba de Hipótesis

pone el énfasis en
intentar detectar la presencia de
un afecto significativo
(grupos que difieren, variables que correlacionan, etc.),

ESTADÍSTICOS MUESTRALES

PUEDEN EXISTIR "N" MUESTRAS DE UNA POBLACIÓN, POR TANTO LA MUESTRA ES UNA VARIABLE ALEATORIA CON FUNCIÓN DE DENSIDAD



POBLACIÓN VS. MUESTRA

PARAMETRO		ESTADÍSTICO
μ	Media	$\bar{X}$
σ^2	Varianza	S^2
σ	Desviación Estándar	S
P	Proporción	p

Estadísticos muestrales. La población tiene parámetros (μ, σ^2) mientras que la muestra tiene estadísticos que son variables aleatorias

ESTADÍSTICOS MUESTRALES

1. **Media muestral. Es la media de la muestra (es una variable aleatoria)**

$$\overline{x} = \frac{\sum x_i}{n}$$

Valores de la muestra

Tamaño de la muestra

$E(\overline{x}) = \mu$ El valor esperado de la media muestral es la media poblacional

$Var(\overline{x}) = \frac{\sigma^2}{n}$ Es la varianza de la media muestral (Raíz cuadrada es el error típico

ESTADÍSTICOS MUESTRALES

2. **Varianza muestral. Es la varianza de la muestra**

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

DISTRIBUCIONES DE ESTADÍSTICOS MUESTRALES

Media muestral (cuando se conocen la varianza poblacional) se distribuye como:

$$\bar{x} \sim N(\mu_x, \frac{\sigma_x^2}{n_x})$$

DISTRIBUCIONES DE ESTADÍSTICOS MUESTRALES

Media muestral (cuando se conocen la varianza poblacional) se distribuye como:

$$\bar{x} \sim t$$

$$con\ n - 1\ grados\ libertad$$

DISTRIBUCIONES DE ESTADÍSTICOS MUESTRALES

Varianza muestral se distribuye como:

$$\frac{nS^2}{\sigma^2} \sim \chi^2$$

$$con\ n - 1\ grados\ libertad$$

DISTRIBUCIONES DE ESTADÍSTICOS MUESTRALES

La proporción se distribuye como:

$$\hat{p} \sim N(p, \frac{pq}{n})$$

DISTRIBUCIONES DE ESTADÍSTICOS MUESTRALES

La diferencia de medias muestrales (cuando se conocen las varianza poblacionales) se distribuye como:

$$\bar{x} - \bar{y} \sim N(\mu_x - \mu_y, \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y})$$

DISTRIBUCIONES DE ESTADÍSTICOS MUESTRALES

La diferencia de medias muestrales (cuando se desconocen las varianza poblacionales y diferentes) se distribuye como:

$$\bar{x} - \bar{y} \sim N(\mu_x - \mu_y, \frac{s_x^2}{n_x - 1} + \frac{s_y^2}{n_y - 1})$$

DISTRIBUCIONES DE ESTADÍSTICOS MUESTRALES

La diferencia de medias muestrales (cuando se desconocen las varianza poblacionales e iguales) se distribuye como:

$$\frac{(\bar{x}-\bar{y})}{\sqrt{\frac{n_x+n_y}{n_x n_y}\frac{n_x s_x^2+n_y s_y^2}{n_x+n_y-2}}} \sim t(n_x+n_y-2)$$

DISTRIBUCIONES DE ESTADÍSTICOS MUESTRALES

Diferencia de proporciones se distribuye como:

$$\widehat{p_x} - \widehat{p_y} \sim N(p_x - p_y, \frac{p_x q_x}{n_x} + \frac{p_y q_y}{n_y})$$

DISTRIBUCIONES DE ESTADÍSTICOS MUESTRALES

Diferencia de varianzas muestrales se distribuye como:

$$S_x^2 - S_y^2 \sim F(m, n) \; donde \; F = \frac{S_x^2}{S_y^2}$$

$$\text{y los grados de libertad son m} - 1 \text{ para la muestra x y n} - 1 \text{ para y}$$

PROPIEDADES DE LOS ESTADÍSTICOS

INSESGADO

Una media de muestra es un estimador no sesgado de la media de la población porque la media de la distribución de muestreo de las medias de muestra tomadas de una población es igual a la media de la población misma.

EL VALOR ESPERADO DEL ESTADÍSTICO ES IGUAL AL VALOR DEL PARÁMETRO POBLACIONAL

PROPIEDADES DE LOS ESTADÍSTICOS

EFICIENTE

Se refiere al tamaño del error estándar de la estadística.

Si comparamos dos estadísticas de una muestra del mismo tamaño y tratamos de decidir cual de ellas es un mejor estimador más eficiente, escogeríamos la que tuviera el menor error estándar o la menor desviación estándar de la distribución de muestreo.

ESTIMADORES INSESGADOS EFICIENTES O DE MÍNIMA VARIANZA

PROPIEDADES DE LOS ESTADÍSTICOS

CONSISTENTE

Si al aumentar el tamaño de la muestra, se tiene la certeza de que el valor del estadístico se aproxima bastante al valor del parámetro de la población.

SUFICIENTE

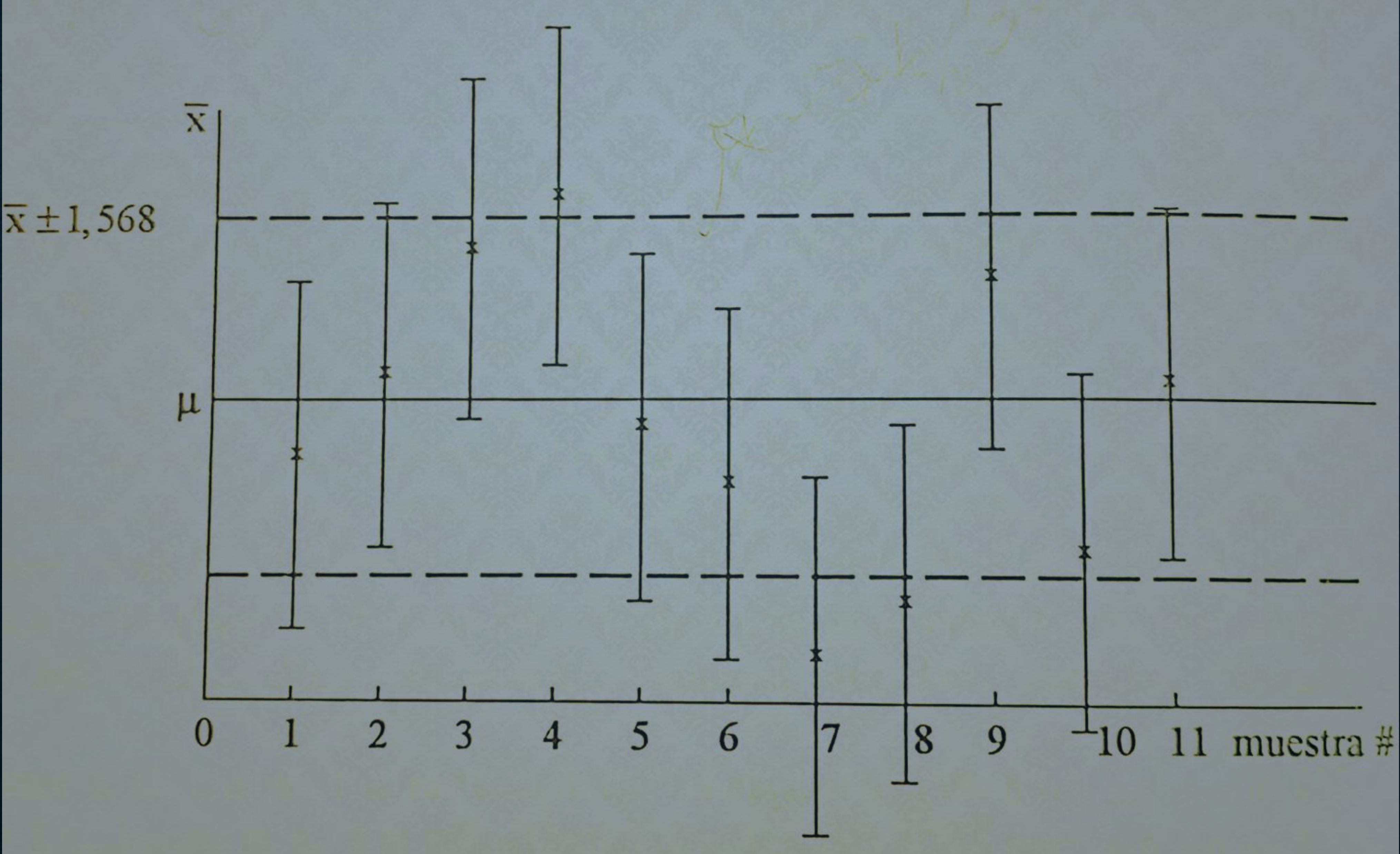
Utiliza una cantidad de la información contenida en la muestra que ningún otro estimador podría extraer información adicional de la muestra sobre el parámetro de la población que se está estimando.

INTERVALOS DE CONFIANZA

INTERVALO DE CONFIANZA

Permite acotar un par o varios pares de valores, dentro de los cuales se encontrará el parámetro poblacional buscada (con una determinada probabilidad).

Un intervalo de confianza nos va a permitir calcular dos valores alrededor de una media muestral (uno superior y otro inferior). Estos valores van a acotar un rango dentro del cual, con una determinada probabilidad, se va a localizar el parámetro poblacional.


$\overline{x}$
$\overline{x} \pm 1,568$
μ
0
1
2
3
4
5
6
7
8
9
10
11
muestra #

INTERVALO DE CONFIANZA

EN TÉRMINOS GENERALES EL INTERVALO SE CALCULA COMO:

$$\bar{x} \pm k_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}$$

INTERVALOS PARA DIFERENCIA DE MEDIAS

$$\bar{x} - \bar{y} \sim N(\mu_x - \mu_y, \frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y})$$

$$\overline{(x} - \overline{y}) \pm k_{\propto/2}\sqrt{\frac{\sigma x^2}{nx} + \frac{\sigma y^2}{ny}}$$

CONTRASTE DE HIPÓTESIS

CONTRASTE DE HIPÓTESIS

- Elemento de toma de decisiones en condiciones de incertidumbre
- El procedimiento para aceptar o rechazar una hipótesis es un contraste de hipótesis

TIPOS DE HIPÓTESIS

- Hipótesis nula (Ho). Sobre la que se toma la decisión. Siempre tiene al menos el signo de =

- Hipótesis alternativa (H1). Es la otra posibilidad

- Hipótesis simple. Si contiene un solo valor (=)

- Hipótesis compuesta. Si contiene más de un valor

($\neq$, $\leq$, $\geq$)

TIPOS DE HIPÓTESIS

Ho		H1	
Simple	Ho: $\theta=\theta o$	Simple	H1: $\theta=\theta 1$
Simple	Ho: $\theta=\theta o$	Compuesta	H1: $\theta>\theta o$ H1: $\theta<\theta o$ H1: $\theta\neq\theta o$
Compuesta	Ho: $\theta\geq\theta o$ Ho: $\theta\leq\theta o$	Compuesta	H1: $\theta>\theta o$ H1: $\theta<\theta o$ H1: $\theta\neq\theta o$
Compuesta		Simple (No es posible porque Ho contiene el signo =)	

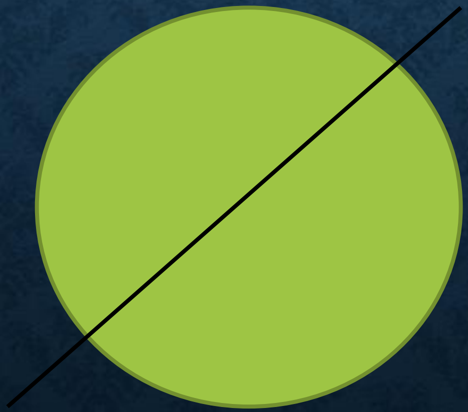
FUNCIÓN O ESTRATEGIA DE DECISIÓN

- Es la función que permite decidir cuando se acepta o se rechaza la Ho

- En el ejemplo anterior se rechaza Ho si el número de accidentes es mayor a 1, es decir, $\bar{x} > 1$

- Divide el espacio muestral en dos subconjuntos:

Región de rechazo o crítica
Permite rechazar la Ho



Región de aceptación
Permite aceptar la Ho

CONTRASTE DE HIPÓTESIS

Es una función de decisión o estrategia que lleva a aceptar o rechazar Ho.

Como la decisión se basa en los datos de una muestra, se pueden cometer dos errores:

- Error tipo I: Rechazar la Ho siendo verdadera.

$$P(I) = \alpha$$

- Error tipo II: Aceptar la Ho siendo falsa.

$$P(II) = \beta$$

ELECCIÓN DEL CONTRASTE

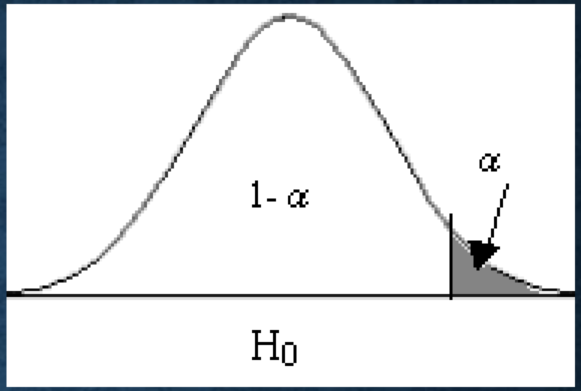
Pueden existir muchas funciones de decisión (contraste de hipótesis), pero debe elegirse aquella que minimizar los errores tipo I y II

Existen entonces tres casos:

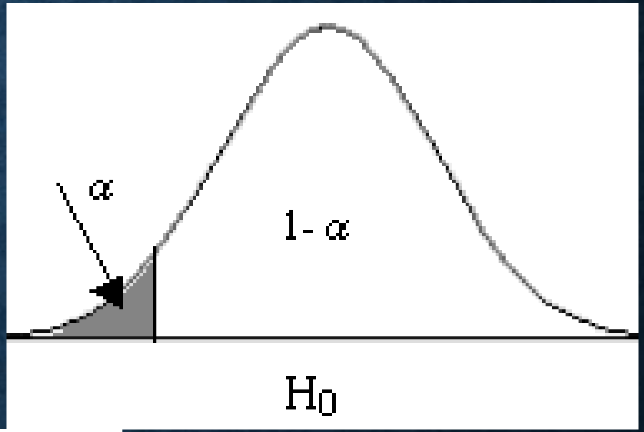
- Que la región crítica de rechazo quede al lado derecho (contraste de una cola por la derecha)
- Que la región crítica o de rechazo quede al lado izquierdo (contraste de una cola izquierda)
- Que la región crítica o de rechazo esté en ambos lados (contraste de dos colas)

ELECCIÓN DEL CONTRASTE

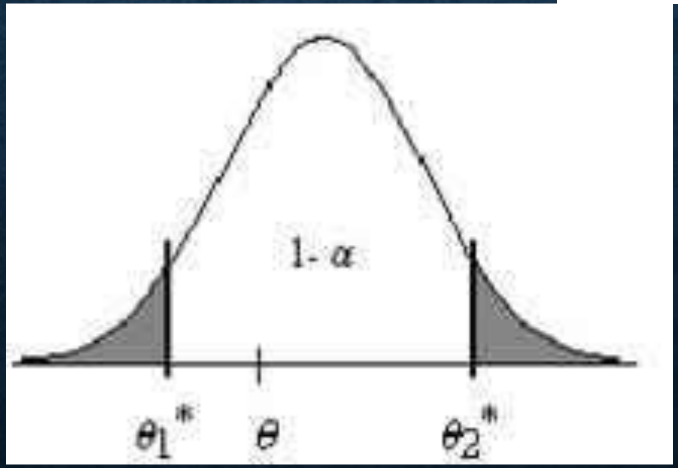
- Contraste de una cola por la derecha



- Contraste de una cola izquierda



- Contraste de dos colas



ELECCIÓN DEL CONTRASTE

Ho	**H1**	**Tipo de contraste**	**Valores críticos**	**Región crítica**
Ho: $\theta=\theta o$	H1: $\theta>\theta o$	Cola derecha	h	Valor estadístico mayor a h
Ho: $\theta=\theta o$	H1: $\theta<\theta o$	Cola izquierda	h	Valor estadístico menor a h
Ho: $\theta=\theta o$	H1: $\theta\neq\theta o$	Dos colas	h1 y h2	Valor estadístico menor a h1 y mayores a h2

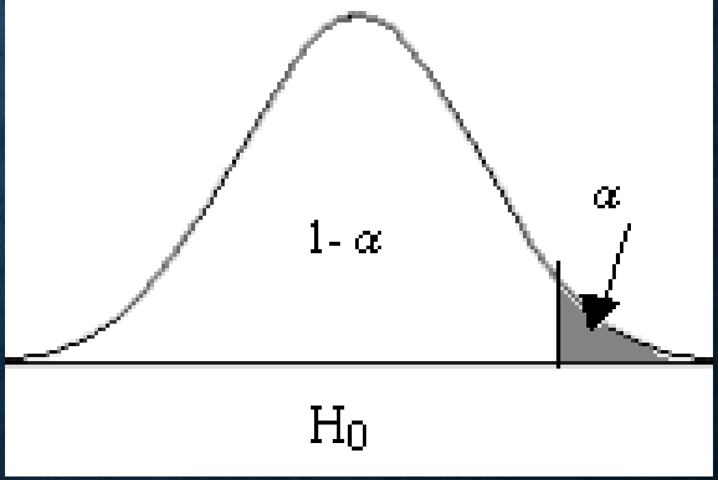
CONTRASTE DE HIPÓTESIS PARAMÉTRICO

CONTRASTE PARA MEDIA DE LA POBLACIÓN / VARIANZA CONOCIDA – COLA A LA DERECHA

$$Ho\colon\ \mu = \mu_o \ \ o \ \ \mu \le \mu_o$$

$$H_1\colon\ \mu > \mu_o \ \ o \ \ \mu = \mu_1\,, \mu_1 > \mu_o$$

$$P\left[\frac{(\bar{x}-\mu)}{\sqrt{\frac{\sigma^2}{n}}} > K_\alpha\right] = \alpha$$



CONTRASTE PARA MEDIA DE LA POBLACIÓN / VARIANZA CONOCIDA – COLA A LA IZQUIERDA

$$Ho\colon \mu = \mu_o \ \ o \ \ \mu \geq \mu_o$$

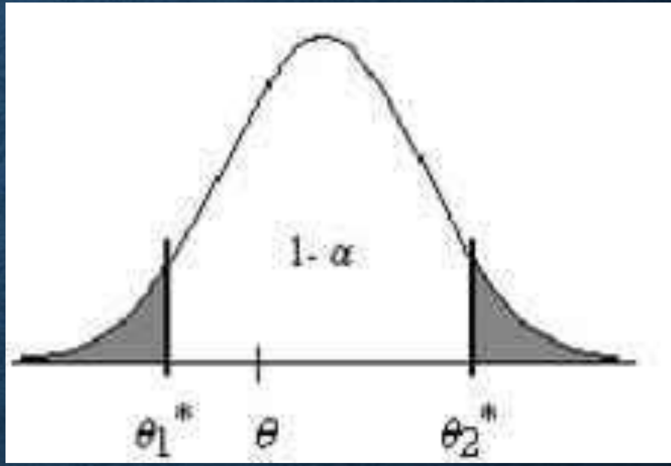
$$H_1\colon \mu < \mu_o \ \ o \ \ \mu = \mu_1 \, , \mu_1 < \mu_o$$

$$P\left[\frac{(\bar{x} - \mu)}{\sqrt{\frac{\sigma^2}{n}}} < K_\alpha\right] = \alpha \qquad Ho\colon \mu = \mu_o \ \ o \ \ \mu \geq \mu_o$$

CONTRASTE PARA MEDIA DE LA POBLACIÓN / VARIANZA CONOCIDA – DOS COLAS

$$Ho: \mu = \mu_o$$

$$H_1: \mu \neq \mu_o$$



$$P\left[-k_{\alpha/2}\frac{(\bar{x}-\mu)}{\sqrt{\frac{\sigma^2}{n}}} < k_{\alpha/2}\right] = 1-\alpha$$

CONTRASTE PARA MEDIA DE LA POBLACIÓN / VARIANZA CONOCIDA – COLA A LA DERECHA

Se toma una muestra de 100 automóviles que transitaron por una calle adoquinada, se produjeron 2 accidentes. Si la desviación típica es 4. Usando un nivel de significancia del 5%. Se debe adoquinar la zona, si el número de accidente promedio es 1.

$Ho\colon \mu = 1 \quad H_1\colon \mu > 1$

$$P\left[\frac{(\bar{x}-\mu)}{\sqrt{\frac{\sigma^2}{n}}} > K_\alpha\right] = \alpha \longrightarrow P\left[\bar{x} > \mu + K_\alpha\sqrt{\frac{\sigma^2}{n}}\right] = \alpha$$

$$\boxed{h = \mu + K_\alpha\sqrt{\frac{\sigma^2}{n}}} \longrightarrow h = 1 + 1{,}645\sqrt{\frac{16}{100}} = 1{,}6458$$

$$\bar{x} > 1{,}6458 \quad Se\ rechaza$$

CONTRASTE PARA MEDIA DE LA POBLACIÓN / VARIANZA CONOCIDA – COLA A LA IZQUIERDA

Se aplicó un programa de refuerzo escolar para mejorar las calificaciones, se espere que la media mejore y llegue a 8,50. Para conocer si efectivamente se mejoraron las calificaciones se aplica una prueba a una muestra de 100 estudiantes, obteniendo una calificación de 8,25, con una desviación típica de la población de 0,30. Contrastar la hipótesis del promedio de 8,50 con un nivel de significancia de 2,5%

$$Ho\colon \mu = 8{,}50 \quad H_1\colon \mu < 8{,}50$$

CONTRASTE PARA MEDIA DE LA POBLACIÓN / VARIANZA CONOCIDA – COLA A LA IZQUIERDA

$$Ho\colon \mu = 8{,}50 \quad H_1\colon \mu < 8{,}50$$

$$P\left[\frac{(\bar{x}-\mu)}{\sqrt{\frac{\sigma^2}{n}}} < -K_\alpha\right] = \alpha \longrightarrow P\left[\bar{x} < \mu - K_\alpha\sqrt{\frac{\sigma^2}{n}}\right] = \alpha$$

$$\boxed{h = \mu - K_\alpha\sqrt{\frac{\sigma^2}{n}}} \longrightarrow h = 8{,}50 - 1{,}96\frac{0{,}30}{\sqrt{100}} = 8{,}44$$

$$\bar{x} < 8{,}44 \quad Se\ rechaza$$

CONTRASTE PARA MEDIA DE LA POBLACIÓN / VARIANZA CONOCIDA – DOS COLAS

Se sabe que el peso promedio de las mujeres en una comunidad es de 72Kg, se aplica una campaña de comida saludable. Se toma una muestra de 16 mujeres y se obtiene un peso promedio de 71,50Kg. La campaña produjo un cambio en el peso? Se sabe que la desviación típica es de 0,250 y se desea contrastar con un nivel de significación de 5%.

$$Ho: \mu = 72 \;\; H_1: \mu \neq 72$$

CONTRASTE PARA MEDIA DE LA POBLACIÓN / VARIANZA CONOCIDA – DOS COLAS

$$Ho: \mu = 72 \;\; H_1: \mu \neq 72$$

$$P\left[-K_\alpha < \frac{(\bar{x}-\mu)}{\sqrt{\frac{\sigma^2}{n}}} < -K_\alpha\right] = 1-\alpha$$

$$P\left[\mu - K_\alpha\sqrt{\frac{\sigma^2}{n}} < \bar{x} < \mu + K_\alpha\sqrt{\frac{\sigma^2}{n}}\right] = 1-\alpha$$

$$\boxed{h_1 = \mu - K_\alpha\sqrt{\frac{\sigma^2}{n}}} \longrightarrow h_1 = 72 - 1{,}96\frac{0{,}250}{\sqrt{16}} = 71{,}87$$

$$\boxed{h_2 = \mu + K_\alpha\sqrt{\frac{\sigma^2}{n}}} \longrightarrow h_2 = 72 + 1{,}96\frac{0{,}250}{\sqrt{16}} = 72{,}1225$$

$$\bar{x} < 71{,}87 \quad Se\ rechaza$$

CONTRASTE PARA PROPORCIONES

UNA COLA A LA DERECHA

$$Ho\!: p = p_o \;\; o \;\; p \leq p_o$$

$$H_1\!: p > p \;\; o \;\; p = p_1\,, p_1 > p_o$$

$$\boxed{h = p_o + K_\alpha \sqrt{\frac{p_o(1-p_o)}{n}}}$$

$Si\ \hat{p} > h \;\; Se\ rechaza$

CONTRASTE PARA PROPORCIÓN

UNA COLA A LA IZQUIERDA

$$Ho: p = p_o \ \ o \ \ p \geq p_o$$

$$H_1: p < p_o \ \ o \ \ p = p_1 \, , p_1 < p_o$$

$$\boxed{h = p_o - K_\alpha \sqrt{\frac{p_o(1-p_o)}{n}}}$$

$Si \ \hat{p} < h \ \ Se \ rechaza$

CONTRASTE PARA LA PROPORCIÓN

DOS COLAS

$$Ho: p = p_o$$

$$H_1: p \neq p_o$$

$$\boxed{h_1 = p_o - K_\alpha \sqrt{\frac{p_o(1-p_o)}{n}}}$$

$$\boxed{h_2 = p_o + K_\alpha \sqrt{\frac{p_o(1-p_o)}{n}}}$$

$$Si\ \hat{p} < h_1\ o\quad \hat{p} > h_2\quad Se\ rechaza$$

RESUMEN – UNA MUESTRA

Tipo contraste	Distribución	Ho	H1	Cola	Regla de decisión	STATA
Media con varianza conocida	Normal cuando se conoce varianza	$\mu = \mu_o$	$\mu > \mu_o$	Cola a la derecha	$\bar{x} < h$ Para aceptar Ho p-value es mayor a 0,10	Ztest (para una muestra)
	Normal cuando se conoce varianza	$\mu = \mu_o$	$\mu < \mu_o$	Cola a la izquierda	$\bar{x} > h$ Para aceptar Ho p-value es mayor a 0,10	Ztest (para una muestra)
	Normal cuando se conoce varianza	$\mu = \mu_o$	$\mu \neq \mu_o$	Dos colas	$\bar{x} < h2$ $\bar{x} > h1$ Para aceptar Ho p-value es mayor a 0,10	Ztest (para una muestra)

Tipo contraste	Distribución	Ho	H1	Cola	Regla de decisión	STATA
Media con varianza desconocida	T-student para varianza desconocida	$\mu = \mu_o$	$\mu > \mu_o$	Cola a la derecha	$\bar{x} < h$ Para aceptar Ho p-value es mayor a 0,10	ttest (para una muestra)
		$\mu = \mu_o$	$\mu < \mu_o$	Cola a la izquierda	$\bar{x} > h$ Para aceptar Ho p-value es mayor a 0,10	ttest (para una muestra)
		$\mu = \mu_o$	$\mu \neq \mu_o$	Dos colas	$\bar{x} < h2$ $\bar{x} > h1$ Para aceptar Ho p-value es mayor a 0,10	ttest (para una muestra)

Tipo contraste	Distribución	Ho	H1	Cola	Regla de decisión	STATA
Proporción	Normal	$\mathrm{p} = p_o$	$\mathrm{p} > p_o$	Cola a la derecha	$\hat{p} < h$ Para aceptar Ho p-value es mayor a 0,10	prtest (para una muestra)
	Normal	$\mathrm{p} = p_o$	$\mathrm{p} < p_o$	Cola a la izquierda	$\hat{p} > h$ Para aceptar Ho p-value es mayor a 0,10	prtest (para una muestra)
	Normal	$\mathrm{p} = p_o$	$\mathrm{p} \neq p_o$	Dos colas	$\hat{p} < h2$ $\hat{\mathrm{p}} > h1$ Para aceptar Ho p-value es mayor a 0,10	prtest (para una muestra)

Tipo contraste	Distribución	Ho	H1	Cola	Regla de decisión	STATA
Varianza	Chi cuadrado	$\sigma^2 = \sigma_o^2$	$\sigma^2 > \sigma_o^2$	Cola a la derecha	$S^2 < h$ Para aceptar Ho p-value es mayor a 0,10	sdtest (para una muestra)
		$\sigma^2 = \sigma_o^2$	$\sigma^2 < \sigma_o^2$	Cola a la izquierda	$S^2 > h$ Para aceptar Ho p-value es mayor a 0,10	sdtest (para una muestra)
		$\sigma^2 = \sigma_o^2$	$\sigma^2 \neq \sigma_o^2$	Dos colas	$S^2 < h2$ $S^2 > h1$ Para aceptar Ho p-value es mayor a 0,10	sdtest (para una muestra)

CONTRASTES PARA DOS MUESTRAS

RESUMEN – DOS MUESTRAS

Tipo contraste	Distribución	Ho	H1	Cola	Regla de decisión	STATA
Diferencia de medias con <u>varianza conocida y muestras independientes</u>	Normal cuando se conoce varianza	$\mu_x = \mu_y$	$\mu_x > \mu_y$	Cola a la derecha	$\bar{x} - \bar{y} < h$ Para aceptar Ho p-value es mayor a 0,10	Ztest (para dos muestras)
	Normal cuando se conoce varianza	$\mu_x = \mu_y$	$\mu_x < \mu_y$	Cola a la izquierda	$\bar{x} - \bar{y} > h$ Para aceptar Ho p-value es mayor a 0,10	Ztest (para dos muestras)
	Normal cuando se conoce varianza	$\mu_x = \mu_y$	$\mu_x \neq \mu_y$	Dos colas	$\bar{x} - \bar{y} < h2$ $\bar{x} - \bar{y} > h1$ Para aceptar Ho p-value es mayor a 0,10	Ztest (para dos muestras)

RESUMEN – DOS MUESTRAS

Tipo contraste	Distribución	Ho	H1	Cola	Regla de decisión	STATA
Diferencia de medias con <u>varianza conocida y muestras dependientes (emparejamiento)</u>	Normal cuando se conoce varianza	$\mu_x = \mu_y$	$\mu_x > \mu_y$	Cola a la derecha	$\bar{x} - \bar{y} < h$ Para aceptar Ho p-value es mayor a 0,10	Ztest (para muestras emparejadas)
	Normal cuando se conoce varianza	$\mu_x = \mu_y$	$\mu_x < \mu_y$	Cola a la izquierda	$\bar{x} - \bar{y} > h$ Para aceptar Ho p-value es mayor a 0,10	Ztest (para muestras emparejadas)
	Normal cuando se conoce varianza	$\mu_x = \mu_y$	$\mu_x \neq \mu_y$	Dos colas	$\bar{x} - \bar{y} < h2$ $\bar{x} - \bar{y} > h1$ Para aceptar Ho p-value es mayor a 0,10	Ztest (para muestras emparejadas)

RESUMEN – DOS MUESTRAS

Tipo contraste	Distribución	Ho	H1	Cola	Regla de decisión	STATA
Diferencia de medias con varianza desconocidas pero diferentes	T-student para varianzas Desconocidas diferentes	$\mu_x = \mu_y$	$\mu_x > \mu_y$	Cola a la derecha	$\bar{x} - \bar{y} < h$ Para aceptar Ho p-value es mayor a 0,10	ttest (para dos muestras y debe indicarse que la varianza es diferente)
(**esto se comprueba con el test de varianza-homocedast.)	T-student para varianzas Desconocidas diferentes	$\mu_x = \mu_y$	$\mu_x < \mu_y$	Cola a la izquierda	$\bar{x} - \bar{y} > h$ Para aceptar Ho p-value es mayor a 0,10	ttest (para dos muestras y debe indicarse que la varianza es diferente)
Para muestras independientes	T-student para varianzas Desconocidas diferentes	$\mu_x = \mu_y$	$\mu_x \neq \mu_y$	Dos colas	$\bar{x} - \bar{y} < h2$ $\bar{x} - \bar{y} > h1$ Para aceptar Ho p-value es mayor a 0,10	ttest (para dos muestras y debe indicarse que la varianza es diferente)

RESUMEN – DOS MUESTRAS

Tipo contraste	Distribución	Ho	H1	Cola	Regla de decisión	STATA
Diferencia de medias con varianza desconocidas iguales	T-student para varianzas Desconocidas diferentes	$\mu_x = \mu_y$	$\mu_x > \mu_y$	Cola a la derecha	$\bar{x} - \bar{y} < h$ Para aceptar Ho p-value es mayor a 0,10	ttest (para dos muestras y no debe indicarse que la varianza es diferente)
(**esto se comprueba con el test de varianza-homocedast.)	T-student para varianzas Desconocidas diferentes	$\mu_x = \mu_y$	$\mu_x < \mu_y$	Cola a la izquierda	$\bar{x} - \bar{y} > h$ Para aceptar Ho p-value es mayor a 0,10	ttest (para dos muestras y no debe indicarse que la varianza es diferente)
Para muestras independientes	T-student para varianzas Desconocidas diferentes	$\mu_x = \mu_y$	$\mu_x \neq \mu_y$	Dos colas	$\bar{x} - \bar{y} < h2$ $\bar{x} - \bar{y} > h1$ Para aceptar Ho p-value es mayor a 0,10	ttest (para dos muestras y no debe indicarse que la varianza es diferente)

RESUMEN – DOS MUESTRAS

Tipo contraste	Distribución	Ho	H1	Cola	Regla de decisión	STATA
Diferencia de medias con <u>varianza desconocidas</u> (**esto se comprueba con el test de varianza-homocedast.) <u>Para muestras emparejadas (dependientes)</u>	T-student para varianzas Desconocidas diferentes	$\mu_x = \mu_y$	$\mu_x > \mu_y$	Cola a la derecha	$\bar{x} - \bar{y} < h$ Para aceptar Ho p-value es mayor a 0,10	ttest (para emparejados)
	T-student para varianzas Desconocidas diferentes	$\mu_x = \mu_y$	$\mu_x < \mu_y$	Cola a la izquierda	$\bar{x} - \bar{y} > h$ Para aceptar Ho p-value es mayor a 0,10	ttest (para emparejados)
	T-student para varianzas Desconocidas diferentes	$\mu_x = \mu_y$	$\mu_x \neq \mu_y$	Dos colas	$\bar{x} - \bar{y} < h2$ $\bar{x} - \bar{y} > h1$ Para aceptar Ho p-value es mayor a 0,10	ttest (para emparejados)

RESUMEN – DOS MUESTRAS

Tipo contraste	Distribución	Ho	H1	Cola	Regla de decisión	STATA
Diferencias de Proporciones	Normal	$p_x = p_y$	$p_x > p_y$	Cola a la derecha	$\hat{p}_x - \hat{p}_y < h$ Para aceptar Ho p-value es mayor a 0,10	prtest (para dos muestras)
	Normal	$p_x = p_y$	$p_x < p_y$	Cola a la izquierda	$\hat{p}_x - \hat{p}_y > h$ Para aceptar Ho p-value es mayor a 0,10	prtest (para dos muestras)
	Normal	$p_x = p_y$	$p_x \neq p_y$	Dos colas	$\hat{p}_x - \hat{p}_y < h2$ $\hat{p}_x - \hat{p}_y > h1$ Para aceptar Ho p-value es mayor a 0,10	prtest (para dos muestras)

RESUMEN – DOS MUESTRAS

Tipo contraste	Distribución	Ho	H1	Cola	Regla de decisión	STATA
** Diferencia de Varianzas	Distribución F	$\sigma_x^2 = \sigma_y^2$	$\sigma_x^2 > \sigma_y^2$	Cola a la derecha	$S_x^2 - S_y^2 < h$ Para aceptar Ho p-value es mayor a 0,10	sdtest (para dos muestras)
	Distribución F	$\sigma_x^2 = \sigma_y^2$	$\sigma_x^2 < \sigma_y^2$	Cola a la izquierda	$S_x^2 - S_y^2 > h$ Para aceptar Ho p-value es mayor a 0,10	sdtest (para dos muestras)
	Distribución F	$\sigma_x^2 = \sigma_y^2$	$\sigma_x^2 \neq \sigma_y^2$	Dos colas	$S_x^2 - S_y^2 < h2$ $S_x^2 - S_y^2 > h1$ Para aceptar Ho p-value es mayor a 0,10	sdtest (para dos muestras)

CONTRASTES NO PARAMÉTRICOS PARA DISTRIBUCIÓN

DETERMINAR SI LA FUNCIÓN DE DISTRIBUCIÓN TEÓRICA ES LA VERDADERA:

Bondad del ajuste de Chi-cuadrado y Kolmorogov – Smirnov

Test	Utilidad	Ho / H1	Regla decisión	STATA
Bondad del ajuste Chi-cuadrado	Determinar si una variable categórica sigue o no una distribución hipotética.	Ho: La distribución es la que se cree H1: No es la que se cree	Si p-value es mayor a 0,10 acepto Ho	Cargar el paquete csgof Comando: csgof race, expperc (70, 20, 10) Emplear prueba sysuse nlsw88
Kolmorogov - Smirnov	Si los datos provienen de una función de distribución determinada. Variables continuas únicamente	Ho: la distribución es normal H1: la distribución no es normal	Si p-value es mayor a 0,10 acepto Ho	ksmirnov x = normal((x-r(mean))/r(sd))