FORMULARIO ESTADÍSTICA II

DISTRIBUCIONES DE MUESTREO

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{x}}}$$

Si n/N≥.05, entonces $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$

ESTIMACIONES PUNTUALES

$$\bar{x} = \frac{\sum xi}{n} \qquad s^2 = \frac{\sum\left(x_i - \bar{x}^2\right)}{n-1} \qquad s = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}}$$

$$\bar{p} = \frac{\#\, exitos\ en\ la\ muestra}{n}$$

ESTIMACIONES DE INTERVALO

Para muestras grandes y σ conocida

$$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}} = \left[\bar{x} - z_{\alpha/2}\sigma_{\bar{x}}, \bar{x} + z_{\alpha/2}\sigma_{\bar{x}}\right]$$

Para muestras grandes y σ desconocida

$S_{\bar{x}} = \frac{s}{\sqrt{n}}$ para poblaciones infinitas

$s_{\bar{x}} = \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ para poblaciones finitas y n/N≥.05

$$\bar{x} \pm z_{\alpha/2} s_{\bar{x}} = \left[\bar{x} - z_{\alpha/2} s_{\bar{x}}, \bar{x} + z_{\alpha/2} s_{\bar{x}}\right]$$

Para muestras pequeñas (n≤30) y σ desconocida

$$\bar{x} \pm t_{\alpha/2} s_{\bar{x}} = \left[\bar{x} - t_{\alpha/2} s_{\bar{x}}, \bar{x} + t_{t/2} s_{\bar{x}}\right]$$

Para porciones y muestras grandes

$$\sigma_{\bar{p}} = \sqrt{\frac{\overline{pq}}{n}}$$

$$\bar{p} \pm z_{\alpha/2} \sigma_{\bar{p}} = \left[\bar{p} - z_{\alpha/2} \sigma_{\bar{p}}, \bar{p} + z_{\alpha/2} \sigma_{\bar{p}}\right]$$

Determinación del tamaño de muestra en la estimación

$$n = \left(\frac{z_{\alpha/2}\sigma}{k}\right)^2$$ para medias

$$n = \frac{z_{\alpha/2}{}^2 pq}{k^2}$$ para porciones

PRUEBA DE HIPÓTESIS (Una muestra)

Prueba de hipótesis cuando se conoce σ.

$$z = \frac{\bar{x} - \mu_{H_0}}{\sigma_{\bar{x}}}$$

- Prueba de 2 extremos:

$H_0: \mu = \mu_{H_0}$
$H_a: \mu \neq \mu_{H_0}$
No rechazar H_0 si $|z| \leq z_{\alpha/2}$

- Prueba de 1 extremo (inferior)

$H_0: \mu \geq \mu_{H_0}$
$H_a: \mu < \mu_{H_0}$
No rechazar H_0 si $z \geq z_{\alpha}$

- Prueba de 1 extremo (superior)

$H_0: \mu \leq \mu_{H_0}$
$H_a: \mu > \mu_{H_0}$
No rechazar H_0 $si\ z \leq z_\alpha$

Prueba de hipótesis de porciones: muestras grandes

$$Z = \frac{\bar{p} - p_{H_0}}{\sigma \bar{p}} \qquad \sigma_{\bar{p}} = \sqrt{\frac{p_{H_0} q_{H_0}}{n}}$$

- Prueba de 2 extremos:

$H_{0:}\ p = p_{H_0}$
$H_{a:}\ p \neq p_{H_0}$
No rechazar $H_{0:}$ $si\ |z| \leq z_{\alpha/2}$

- Prueba de 1 extremo (inferior)

$H_{0:}\ p \geq p_{H_0}$
$H_{a:}\ p < p_{H_0}$
No rechazar $H_{0:}$ $si\ z \geq z_\alpha$

- Prueba de 1 extremo (superior)

$H_{0:}\ p \leq p_{H_0}$
$H_{a:}\ p > p_{H_0}$
No rechazar $H_{0:}$ $si\ z \leq z_\alpha$

Prueba de hipótesis de medias cuando no se conoce σ

$$t = \frac{\bar{x} - \mu_{H_0}}{s_{\bar{x}}} \qquad s_{\bar{x}} = \frac{s}{\sqrt{n}}$$

- Prueba de 2 extremos

$H_0: \mu = \mu_{H_0}$
$H_a: \mu \neq \mu_{H_0}$
No rechazar H_0 si $|t| \leq t_{\alpha/2}$

- Prueba de 1 extremo (inferior)

$H_0: \mu \geq \mu_{H_0}$
$H_a: \mu > \mu_{H_0}$
No rechazar H_0 $si\ t \geq t_\alpha$

- Prueba de un extremo (superior)

$H_0: \mu \leq \mu_{H_0}$
$H_a: \mu > \mu_{H_0}$

No rechazar H_0 si $t \leq t_\alpha$

PRUEBA DE HIPÓTESIS (Dos muestras)

Diferencias entre medias: tamaños de muestras grandes

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}} \qquad S_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}} \quad z = \frac{\bar{x}_1-\bar{x}_2-(\mu_1-\mu_2)}{s_{\bar{x}_1-\bar{x}_2}}$$

- Prueba de 2 extremos

$H_0: \mu_1 - \mu_2 = 0$
$H_a: \mu_1 - \mu_2 \neq 0$
No rechazar H_0 si $|z| \leq z_{\alpha/2}$

- Prueba de extremo inferior

$H_0: \mu_1 - \mu_2 \geq 0$
$H_a: \mu_1 - \mu_2 < 0$
No rechazar H_0 si $|z| \geq z_\alpha$

- Prueba de extremo superior

$H_0: \mu_1 - \mu_2 \leq 0$
$H_a: \mu_1 - \mu_2 > 0$
No rechazar H_0 si $z \leq z_\alpha$

Pruebas para diferencias entre medias: tamaños de muestra pequeños

$$s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_1^2}{n_1+n_2-2}; \; s_{\bar{x}_1-\bar{x}_2} = s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}; \quad t = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{s_{\bar{x}_1-\bar{x}_2}}$$

Buscar los valores t con $n_1 + n_2 - 2$ grados de libertad

- Prueba de 2 extremos

$H_0: \mu_1 - \mu_2 = 0$
$H_a: \mu_1 - \mu_2 \neq 0$
No rechazar H_0 si $|t| \leq t_{\alpha/2}$

- Prueba de extremo inferior

$H_0: \mu_1 - \mu_2 \geq 0$
$H_a: \mu_1 - \mu_2 < 0$
No rechazar H_0 si $t \geq t_\alpha$

- Prueba de extremo superior

$H_0: \mu_1 - \mu_2 \leq 0$
$H_\alpha: \mu_1 - \mu_2 > 0$
No rechazar H_0 si $t \leq t_\alpha$

Prueba de diferencias entre medias con muestras dependientes

$$t = \frac{\bar{x}_d - \mu_d}{s_{\bar{d}}} \qquad s_{\bar{d}} = \frac{s_d}{\sqrt{n}}$$

- Prueba de 2 extremos

$H_0: \mu_d = 0$
$H_\alpha: \mu_d \neq 0$
No rechazar H_0 si $|t| < t_{\alpha/2}$
$|t| < t_{\alpha/2}$

- Prueba de extremo inferior

$H_0: \mu_d \geq 0$
$H_\alpha: \mu_d < 0$
No rechazar H_0 si $t \geq t_\alpha$
$t \geq t \propto$

- Prueba de extremo superior

$H_0: \mu_d \leq 0$
$H_\alpha: \mu_d > 0$
No rechazar H_0 si $t \leq t_\alpha$
$t \leq t_\alpha$

Pruebas para diferencias entre porciones: muestras grandes

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} \qquad s_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}} \quad z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)}{s_{\bar{p}_1 - \bar{p}_2}}$$

- Prueba de 2 extremos

$H_0: p_1 - p_2 = 0$
$H_\alpha: p_1 - p_2 \neq 0$
No rechazar si $|z| \leq z_{\alpha/2}$

- Prueba de extremo inferior

$H_0: p_1 - p_2 \geq 0$
$H_\alpha: p_1 - p_2 < 0$
No rechazar si $z \geq z_\alpha$

- Prueba de extremo superior

$H_0: p_1 - p_2 \leq 0$
$H_a: p_1 - p_2 > 0$

No rechazar si $z \leq z_\alpha$

Valor "p"

De 2 extremos
$p = 1 - 2\,p(z)$

De 1 extremo
$p = 0.5 - p(z)$

No rechazar Ho si $p \geq \alpha$

JI CUADRADA COMO PRUEBA DE INTERDEPENDENCIA.

H_0: La variable de columna es independiente de la variable de fila.

H_0: la variable de columna no es independiente de la variable de fila.

-Obtener la tabla de contingencia con las frecuencias observadas (f_i).
-Obtener la tabla de frecuencias esperadas (e_i).

$$e_{ij} = \frac{(Total\ de\ fila)(Total\ de\ columna)}{n}$$

-Obtener $\chi^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{(f_{ij} - e_{ij})^2}{e_{ij}}$

*Buscar χ^2_α con (n-1)(m-1) grados de libertad y α nivel de significancia.

No rechazar H_0 si $\chi^2 \leq \chi^2_\alpha$

PRUEBA DE BONDAD DE AJUSTE : DISTRIBUCIÓN BINOMIAL, POISSON Y NORMAL.

Ho: La población tiene una distribución (binmial, poisson, o normal) de probabilidad.

Ho: La población no tiene una distribución (binomial , poisson o normal) de probabilidad.

-Tomar una muestra aleatoria y asentar las frecuencias observadas (*fi)*
-Calcular las frecuencias esperadas (*ei)*

-Calcular el estadístico
$$\chi^2 = \sum \frac{(f_i - e_i)^2}{e_i}$$
$$e_i = p_i n$$

-No rechazar Ho si $\chi^2 \leq \chi_\alpha^{\ 2}$ para α nivel de significación.

**Nota : para hacer una distribución binomial °L= k-r-1 donde r son las estimaciones puntuales utilizadas.

Distribución binomial $P(r) = {}_n C_r p^r q^{n-r}$

Distribución Poisson. $$P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

ANÁLISIS DE VARIANZA

Ho: $\mu_1 = \mu_2 = \ldots\ldots.\mu_k$
Ho: No todas las medias μ_i son iguales.

-Calcular la varianza población dentro de las muestras $\hat{\sigma} = \sum_{j=1}^{k} \frac{n_j \left(\bar{x}_j - \bar{\bar{x}} \right)^2}{k-1}$ donde:

$$\bar{\bar{x}} = \frac{\sum_{j=1}^{k} \sum_{i=1}^{n_j} x_{ij}}{n_T}$$

x_{ij}= observación i de la muestra j
n_j: tamaño de la muestra j

$n_T = \sum_{j}^{k} n_j$ suma de los tamaños de la muestra.

-Calcular la varianza población dentro de las muestras $\overline{\sigma}^2$

$$\overline{\sigma}^2 = \frac{\sum_{j=1}^{k} (n_j - 1) s_j^2}{n_T - k}$$

donde:

$$s_j^2 = \frac{\sum_{j=1}^{n_j} (n - 1) s_j^2}{n_T - k}$$

$$\overline{x}_j = \frac{\sum_{i=1}^{n_j} \left(x_{ij} - \overline{x}_j \right)^2}{n_j}$$

-Calcular el estadístico de prueba $F = \frac{\hat{\sigma}}{\overline{\sigma}}$

-No rechazar Ho= si $F \leq F_\alpha$

donde F_α se extrae de la distribución F con k-1 grados de libertad en el numerador, y $n_T - k$ en el denominador.

INFERENCIAS ACERCA DE LA VARIANZA DE UNA POBLACIÓN

$\frac{(n-1)s^2}{\chi^2_{\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}}$ intervalo de confianza para la varianza. Donde los valores χ^2 se basan en la distinción ji cuadrada con n-1 grados de libertad y $1-\alpha$ es el coeficiente de confianza.

Prueba de Hipótesis

$$\chi^2 = \frac{(n - 1) s^2}{\sigma^2_{H_0}}$$

-De dos extremos

$$H_0 : \sigma^2 = \sigma_{Ho}{}^2$$
$$H_a : \sigma^2 \neq \sigma_{Ho}{}^2$$

No rechazar Ho si $\chi^2{}_{1-\alpha/2} \leq \chi^2 \leq \chi_{\alpha/2}{}^2$

-De extremo inferior:

$$H_0 : \sigma^2 \geq \sigma_{H_0}{}^2$$
$$H_a : \sigma^2 < \sigma_{H_0}{}^2$$

No rechazar Ho si : $\chi^2 \geq \chi_{1-\alpha}{}^2$

-De extremo superior

$$H_0 : \sigma^2 \leq \sigma_{H_0}{}^2$$
$$H_a : \sigma^2 > \sigma_{H_0}{}^2$$

No rechazar Ho si : $\chi^2 \leq \chi_{1-\alpha}{}^2$

INFERENCIAS ACERCA DE LA VARIANZA DE DOS POBLACIONES

$$F = \frac{s_1{}^2}{s_2{}^2} \qquad F(n_1 - 1, n_2 - 1, 1 - \alpha / 2) = \frac{1}{F(n_1 - 1, n_2 - 1, \alpha / 2)}$$

F_α $y F_{\alpha/2}$ Se basan en la distribución F con $n_1 - 1$ grados de libertad en el numerador y con $n_2 - 1$ grados en el denominador.

-De dos extremos

$$H_0 : \sigma_1{}^2 = \sigma_2{}^2$$
$$H_a : \sigma_1{}^2 \neq \sigma_2{}^2$$

No rechazar Ho si $F \leq F_{\alpha/2}$

-De extremo superior

$$H_0 : \sigma_1^2 \leq \sigma_2^2$$
$$H_a : \sigma_1^2 > \sigma_2^2$$

No rechazar Ho si $F \leq F\alpha$

REGRESIÓN LINEAL

$$\hat{y} = a + bx$$

$$b = \frac{\sum xy - n\overline{X}\overline{Y}}{\sum x^2 - n\overline{x}^2}$$

$$a = \overline{y} - b\overline{X}$$

ERROR ESTÁNDAR DE LA ESTIMACIÓN Y COEFICIENTE DE DETERMINACIÓN

$$s_e = \sqrt{\frac{\sum (y - \hat{y})^2}{n-2}}$$

$$s_e = \sqrt{\frac{\sum y^2 - a\sum y - b\sum xy}{n-2}}$$

$$r^2 = 1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \overline{y})^2}$$

$$r^2 = 1 - \frac{a\sum y + b\sum xy - n\overline{y}^2}{\sum y^2 - n\overline{y}^2}$$ Coeficiente de determinación.

$$r = \sqrt{r^2}$$ Coeficiente de correlación

L a estimación de $\hat{y}(x)$ con un nivel de confianza (con n-2 grados de libertad) es:

$$\hat{y}(x) \pm t_{\alpha/2} s_e$$