RESUMEN DE FÓRMULAS – EXAMEN FINAL

Teorema de Bayes

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A|B)P(B) = P(B|A)P(A) = P(A \cap B)$$

Ley de Probabilidades Totales

$$\Omega = \{N1 \cup N2\} \quad \rightarrow \quad P(A) = P(A \cap N1) + P(A \cap N2)$$

Valor Esperado y Varianza

$$E[X] = \mu = \sum_{R(X)} x_i g_X(x_i) \quad \text{ó} \quad \int_{R(X)} x f_X(x)\, dx$$

$$Var[X] = E[X^2] - (E[X])^2$$

Regla de la Multiplicación

$$\#\, Total\ de\ resultados = n_1 * n_2 * \dots * n_r$$

Muestra de Orden

$$\#m(r; A) = n^r$$

Permutaciones

$$P(r; n) = \frac{n!}{(n-r)!};\ r \leq n$$

Combinaciones

$$C(r; n) = \frac{n!}{(n-r)!\, r!}$$

Particiones Ordenadas

$$\#P(n_1, n_2, \dots, n_r; A) = \binom{N}{n_1}\binom{N-n_1}{n_2}\dots\binom{N-n_1-n_2-\dots-n_{r-1}}{n_r}$$

$$\#P(n_1, n_2, \dots, n_r; A) = \frac{N!}{n_1!\, n_2! \dots n_r!}$$

DISTRIBUCIONES	Parámetros	Función de Probabilidad / fdp	E(X)	Var(X)
DISTRIBUCIÓN UNIFORME DISCRETA	$k =$ número de valores distintos que puede tomar X	$\frac{1}{k}$ $x = x_1, x_2, \dots$	$\frac{1}{k}\sum_{i=1}^{k} x_i$	$\frac{1}{k}\sum_{i=1}^{k}(x_i-\mu)^2$
BERNOULLI	$p =$ probabilidad de éxito	$p^x(1-p)^{1-x}$ $x = 0,1$	p	$p(1-p)$
BINOMIAL	$p =$ probabilidad de éxito $n =$ número de ensayos	$\binom{n}{x} p^x (1-p)^{n-x}$ $0 \leq x \leq n$	np	$np(1-p)$
GEOMÉTRICA	$p =$ probabilidad de éxito	$(1-p)^{x-1}p$ $x = 1,2,\dots$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
BINOMIAL NEGATIVA	$p =$ probabilidad de éxito $k-$ésimo éxito	$\binom{x-1}{k-1}(1-p)^{x-k}p^k$ $x \geq k$	$\frac{k}{p}$	$\frac{k(1-p)}{p^2}$
POISSON	$\lambda = \frac{\#\ de\ arribos}{und.\ tiempo}$ $t =$ intervalo de tiempo	$\frac{(\lambda t)^x}{x!} e^{-\lambda t}$ $x = 0,1,\dots$ $\lambda, t > 0$	λt	λt
DISTRIBUCIÓN UNIFORME CONTINUA	$a =$ mínimo $b =$ máximo	$\frac{1}{b-a}$ $\quad a \leq x \leq b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
DISTRIBUCIÓN NORMAL	$\mu =$ Media $\sigma^2 =$ Varianza	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$	μ	σ^2
	Recuerde: $P(X_N \leq x) = P\left(\frac{X_N-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = P\left(Z \leq \frac{x-\mu}{\sigma}\right)$			
DISTRIBUCIÓN EXPONENCIAL	$\lambda = \frac{\#\ de\ arribos}{und.\ tiempo}$	$\lambda e^{-\lambda x}$ $\quad x, \lambda > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
DISTRIBUCIÓN TRIANGULAR	$a =$ mínimo $m =$ moda $b =$ máximo	$\begin{cases} \frac{2(x-a)}{(m-a)(b-a)} & a \leq x < m \\ \frac{2(b-x)}{(b-m)(b-a)} & m \leq x < b \end{cases}$	$\frac{a+b+m}{3}$	$\frac{a^2+b^2+m^2-ab-am-bm}{18}$

Distribuciones Marginales

- **Discretas**

$$g_X(x) = \sum_{R(Y)} g_{XY}(x,y)$$

- **Continuas**

$$f_X(x) = \int_{R(Y)} f_{XY}(x,y)dy$$

Distribución Condicional

$$f_{X|Y}(x,y) = \frac{f_{XY}(x,y)}{f_Y(y)}$$

Valor Esperado

$$E(u(X,Y)) = \int_{R(Y)}\int_{R(X)} u(X,Y)f_{XY}(X,Y)dx\,dy$$

Varianza y Covarianza de una V.A.

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

$$Var(X) = E[(x-\mu)^2] = \sum_{R(X)}(x-\mu)^2 f_X(x) = \int_{-\infty}^{\infty}(x-\mu)^2 f_X(x)dx$$

Valor Esperado Condicional

$$E(X|Y=b) = \sum_{R(X|Y)} x f_{X|Y}(x,Y=b)$$

$$E(X|Y=b) = \int_{R(X|y=b)} x f_{X|Y}(x,Y=b)dx$$

Probabilidades Condicionales

$$P(X \le a|Y=b) = \int_{R(X\le a|y=b)} f_{X|Y}(x,Y=b)dx$$

$$P(X \le a|Y \le b) = \frac{P(X \le a \cap Y \le b)}{P(Y \le b)}$$

Coeficiente de Correlación

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

Suma de VAs independientes

Sean $X_1\, y\, X_2$ variables aleatorias independientes con función generadoras de momentos $\psi_{x1}(t) y\, \psi_{x2}(t)$,respectivamente, y sea $W = X_1 + X_2$
Entonces:

$$\psi_W(t) = \psi_{x1}(t) * \psi_{x2}(t)$$

Sesgo

$$E(\hat{\theta}) - \theta$$

Mínima Varianza

$$Var(\hat{\theta}_1) < Var(\hat{\theta}_2)$$

Error Cuadrático Medio (ECM)

$$ECM(\hat{\theta}) = Var(\hat{\theta}) + [Sesgo(\hat{\theta})]^2$$

Consistencia

$$\lim_{n\to\infty} E(\hat{\theta}) = \theta\,;\ \lim_{n\to\infty} Var(\hat{\theta}) = 0$$

Cota Rao-Cramer (CRC)

$$CRC = \frac{1}{nE\left(\left[\frac{\delta}{\delta\theta}(\ln f_X(x;\theta))\right]^2\right)} = \frac{1}{-nE\left(\left[\frac{\delta^2}{\delta\theta^2}(\ln f_X(x;\theta))\right]\right)}$$

Estimador de Máxima Verosimilitud

1) $L(x_1,\dots,x_n;\theta) = \prod_{i=1}^{n} f_X(x_i)$
2) $ln(L(x_1,\dots,x_n;\theta))$
3) $\frac{\delta}{\delta\theta}[ln(l)(x_1,\dots,x_n)] = 0$
4) $despejar\ \theta = \hat{\theta}$

Bondad de Ajuste

$$Y_{(m-1)} = \frac{(O_i - E_i)^2}{E_i} = \sum_{i=1}^{m}\frac{(x_i - np_i)^2}{np_i} \sim \chi^2_{(m-1)}$$

$$x_i = Frecuencia\ Observada,\ n = \#\ de\ Observaciones\ Totales$$

$$m = \#\ de\ Clases$$

$$Si\ Y_{calculado} > \chi^2{}_{(m-1;(1-\alpha))} \to rechazo\ H_o$$

Distribuciones Muestrales

χ^2: $\boldsymbol{X_1^2 + X_2^2 + \cdots + X_n^2 \to \chi^2_{(n)}}$
Donde X_i es N(0,1) y X_i es independiente de X_j.

t: $\boldsymbol{\frac{X}{\sqrt{\frac{Y}{n}}} \to t_{(n)}}$
Donde X es N(0,1) y Y es $\chi^2{}_{(n)}$, siendo X y Y independientes.

F: $\boldsymbol{\frac{X/n1}{Y/n2} \to F_{(n1,n2)}}$
Donde X y Y son Vas independientes con distribuciones $\chi^2{}_{(n1)}$ y $\chi^2{}_{(n2)}$ respectivamente.

Si $\boldsymbol{X \to F_{(\alpha;n1.n2)}\ entonces\ Y = \frac{1}{X} \to F_{(1-\alpha;n2.n1)}}$

Intervalos de Confianza

MEDIA POBLACIONAL(μ)			
Distr. Pob.	**Var. Pob.**	**Tamaño Muestra**	**Intervalo de Confianza (1-α)%**
Normal	Conocida	Pequeño	$\bar{X} \pm z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}}$
Normal	Desconocida	Pequeño	$\bar{X} \pm t_{\left(1-\frac{\alpha}{2};n-1\right)} \frac{S}{\sqrt{n}}$
Cualquiera	Conocida	Grande	$\bar{X} \pm z_{\left(1-\frac{\alpha}{2}\right)} \frac{\sigma}{\sqrt{n}}$
Cualquiera	Desconocida	Grande	$\bar{X} \pm z_{\left(1-\frac{\alpha}{2}\right)} \frac{s}{\sqrt{n}}$

ESTIMACIÓN DE UNA PROPORCIÓN POBLACIONAL (p)			
Distr. Pob.	**Var. Pob.**	**Tamaño Muestra**	**Intervalo de Confianza (1- α)%**
Binomial, Bernoulli, Binomial Negativa	Desconocida	Grande	$\hat{p} \pm z_{\left(1-\frac{\alpha}{2}\right)} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $\hat{p} = \frac{x}{n}$, donde $x\ es\ \#\ éxitos$

ESTIMACIÓN DE UNA VARIANZA POBLACIONAL (σ^2)		
Distr. Pob.	**Var. Pob.**	**Intervalo de Confianza (1- α)%**
Normal	Cualquiera	$\left(\frac{(n-1)S^2}{\chi^2_{\left(1-\frac{\alpha}{2};(n-1)\right)}}; \frac{(n-1)S^2}{\chi^2_{\left(\frac{\alpha}{2};(n-1)\right)}}\right)$

ESTIMACIÓN DE UNA DIFERENCIA DE MEDIAS POBLACIONALES $\mu_1-\mu_2$			
Distr. Pob.	**Var. Pob.**	**Tamaño Muestra**	**Intervalo de Confianza (1- α)%**
Normales	Desconocidas e iguales	Pequeños	$\bar{X}_1 - \bar{X}_2 \pm t_{\left(1-\frac{\alpha}{2};(n_1+n_2-2)\right)} S_c \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$ $donde\ S_c = \sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}}$
	Conocidas	Cualquiera	$\bar{X}_1 - \bar{X}_2 \pm z_{\left(1-\frac{\alpha}{2}\right)} \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$
Cualquiera	Conocidas	Grande	$\bar{X}_1 - \bar{X}_2 \pm z_{\left(1-\frac{\alpha}{2}\right)} \sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$

ESTIMACIÓN DE UNA DIFERENCIA DE PROPORCIONES POBLACIONALES p_1-p_2			
Distr. Pob.	**Var. Pob.**	**Tamaño Muestra**	**Intervalo de Confianza (1- α)%**
Bernoulli	Desconocidas	Grandes	$\hat{p}_1 - \hat{p}_2 \pm z_{\left(1-\frac{\alpha}{2}\right)} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

ESTIMACIÓN DE UN COCIENTE DE VARIANZAS POBLACIONALES σ_1^2/σ_2^2		
Distr. Pob.	**Var. Pob.**	**Intervalo de Confianza (1- α)%**
Normales	Cualquiera	$\left(\frac{S_1^2}{S_2^2} F_{\left(\frac{\alpha}{2};(n_2-1);(n_1-1)\right)}; \frac{S_1^2}{S_2^2} F_{\left(1-\frac{\alpha}{2};(n_2-1);(n_1-1)\right)}\right)$

Pruebas de Hipótesis

Prueba de Hipótesis para la Media Poblacional

Hipótesis Nula	Hipótesis Alterna	Supuestos	Estadístico de Prueba (EP)	Región Crítica de Rechazo
$H_0: \mu = \mu_0$	$H_1: \mu < \mu_0$	$X \rightarrow N(\mu, \sigma_0^2)$ $\sigma_0^2: conocida$ $X_1, X_2, \dots, X_n$	$\dfrac{\bar{X} - \mu_0}{\sigma_0/\sqrt{n}} \rightarrow N(0,1)$	$EP < -z_{(1-\alpha)}$
	$H_1: \mu > \mu_0$			$EP > z_{(1-\alpha)}$
	$H_1: \mu \neq \mu_0$			$EP < -z_{(1-\alpha/2)}$ $EP > z_{(1-\alpha/2)}$
$H_0: \mu = \mu_0$	$H_1: \mu < \mu_0$	$X \rightarrow N(\mu, \sigma^2)$ $\sigma^2: desconocida$ $X_1, X_2, \dots, X_n$	$\dfrac{\bar{X} - \mu_0}{S/\sqrt{n}} \rightarrow t_{(n-1)}$	$EP < -t_{(1-\alpha);n-1}$
	$H_1: \mu > \mu_0$			$EP > t_{(1-\alpha);n-1}$
	$H_1: \mu \neq \mu_0$			$EP < -t_{(1-\alpha/2);n-1}$ $EP > t_{(1-\alpha/2);n-1}$

Prueba de Hipótesis para la Diferencia de Medias Poblacionales

Hipótesis Nula	Hipótesis Alterna	Supuestos	Estadístico de Prueba (EP)	Región Crítica de Rechazo
$H_0: \mu_X - \mu_Y = a$	$H_1: \mu_X - \mu_Y < a$	$X \rightarrow N(\mu_X, \sigma_X^2)$ $Y \rightarrow N(\mu_Y, \sigma_Y^2)$ $\sigma_X^2, \sigma_Y^2: conoc$ X y Y independientes $X_1, X_2, \dots, X_{n_X}$ $Y_1, Y_2, \dots, Y_{n_Y}$ *Si las poblaciones no se distribuyen Normal el tamaño de muestra debe ser grande, $n_X, n_Y \geq 3 \rightarrow$ TLC*	$\dfrac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\dfrac{\sigma_X^2}{n_X} + \dfrac{\sigma_Y^2}{n_Y}}} \rightarrow N(0,1)$	$EP < -z_{(1-\alpha)}$
	$H_1: \mu_X - \mu_Y > a$			$EP > z_{(1-\alpha)}$
	$H_1: \mu_X - \mu_Y \neq a$			$EP < -z_{(1-\alpha/2)}$ $EP > z_{(1-\alpha/2)}$
$H_0: \mu_X - \mu_Y = a$	$H_1: \mu_X - \mu_Y < a$	$X \rightarrow N(\mu_X, \sigma^2)$ $Y \rightarrow N(\mu_Y, \sigma^2)$ X y Y independientes $\sigma^2: desconocidas\ e\ iguales$ $X_1, X_2, \dots, X_{n_X}$ $Y_1, Y_2, \dots, Y_{n_Y}$	$\dfrac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{Sp\sqrt{\dfrac{1}{n_X} + \dfrac{1}{n_Y}}} \rightarrow t_{(n_X+n_Y-2)}$ $Sp = \sqrt{\dfrac{(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2}{n_X + n_Y - 2}}$	$EP < -t_{(1-\alpha;\, n_X+n_Y-2)}$
	$H_1: \mu_X - \mu_Y > a$			$EP > t_{(1-\alpha;\, n_X+n_Y-2)}$
	$H_1: \mu_X - \mu_Y \neq a$			$EP < -t_{\left(1-\frac{\alpha}{2};\, n_X+n_Y-2\right)}$ $EP > t_{\left(1-\frac{\alpha}{2};\, n_X+n_Y-2\right)}$

Prueba de Hipótesis para la Proporción

Hipótesis Nula	Hipótesis Alterna	Supuestos	Estadístico de Prueba (EP)	Región Crítica de Rechazo
$H_0: p = p_0$	$H_1: p < p_0$	$X \rightarrow Bernoulli\ (p)$ $\sigma^2: desconocida$ $X_1, X_2, \dots, X_n$ $n \geq 30$ $\hat{p} = \bar{X}$	$\dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1-p_0)}{n}}} \sim N(0,1)$	$EP < -z_{(1-\alpha)}$
	$H_1: p > p_0$			$EP > z_{(1-\alpha)}$
	$H_1: p \neq p_0$			$EP < -z_{(1-\alpha/2)}$ $EP > z_{(1-\alpha/2)}$

Prueba de Hipótesis para la Diferencia de Proporciones

Hipótesis Nula	Hipótesis Alterna	Supuestos	Estadístico de Prueba (EP)	Región Crítica de Rechazo
$H_0: p_x - p_y = 0$	$H_1: p_x - p_y < 0$	$X \rightarrow Bernoulli\ (p_x)$ $Y \rightarrow Bernoulli\ (p_y)$ X y Y in'dependientes $X_1, X_2, \dots, X_{n_X}$ $Y_1, Y_2, \dots, Y_{n_Y}$ $n_X, n_Y \geq 30$ $\widehat{p_X} = \bar{X}\ y\ \widehat{p_Y} = \bar{Y}$	$\dfrac{\hat{p}_x - \hat{p}_y - (p_x - p_y)}{\sqrt{\hat{p}\hat{q}\left(\dfrac{1}{n_x} + \dfrac{1}{n_y}\right)}} \sim N(0,1)$ $\hat{p} = \dfrac{n_x\hat{p}_x + n_y\hat{p}_y}{n_x + n_y}$ $\quad \hat{q} = 1 - \hat{p}$	$EP < -z_{(1-\alpha)}$
	$H_1: p_x - p_y > 0$			$EP > z_{(1-\alpha)}$
	$H_1: p_x - p_y \neq 0$			$EP < -z_{(1-\alpha/2)}$ $EP > z_{(1-\alpha/2)}$

Prueba de Hipótesis para la Varianza

Hipótesis Nula	Hipótesis Alterna	Supuestos	Estadístico de Prueba (EP)	Región Crítica de Rechazo
$H_0: \sigma^2 = {\sigma^2}_0$	$H_1: \sigma^2 < {\sigma^2}_0$	$X \rightarrow N(\mu, \sigma^2)$ $\sigma^2: desconocida$ $X_1, X_2, \dots, X_n$	$\dfrac{(n-1)S^2}{{\sigma^2}_0} \rightarrow \chi^2_{(n-1)}$	$EP < \chi^2_{(\alpha;\, n-1)}$
	$H_1: \sigma^2 > {\sigma^2}_0$			$EP > \chi^2_{(1-\alpha;\, n-1)}$
	$H_1: \sigma^2 \neq {\sigma^2}_0$			$EP < \chi^2_{(\alpha/2;\, n-1)}$ $EP > \chi^2_{(1-\alpha/2;\, n-1)}$

Prueba de Hipótesis para las Varianzas de dos Poblaciones

Hipótesis Nula	Hipótesis Alterna	Supuestos	Estadístico de Prueba (EP)	Región Crítica de Rechazo
$H_0: \dfrac{\sigma_X^2}{\sigma_Y^2} = 1$	$H_1: \dfrac{\sigma_X^2}{\sigma_Y^2} > 1$	$X \rightarrow N(\mu_X, \sigma_X^2)$ $Y \rightarrow N(\mu_Y, \sigma_Y^2)$ $\sigma_X^2\ y\ \sigma_Y^2: desconocidas$ X y Y independientes *X corresponde a la población que tiene la mayor varianza muestral.*	$\dfrac{S_X^2}{S_Y^2} \rightarrow F_{(n_X-1;\, n_Y-1)}$	$EP > F_{(1-\alpha;\, n_X-1;\, n_Y-1)}$
	$H_1: \dfrac{\sigma_X^2}{\sigma_Y^2} < 1$			$EP < F_{(\alpha;\, n_X-1;\, n_Y-1)}$
	$H_1: \dfrac{\sigma_X^2}{\sigma_Y^2} \neq 1$			$EP < F_{(\alpha/2;\, n_X-1;\, n_Y-1)}$ $EP > F_{(1-\alpha/2;\, n_X-1;\, n_Y-1)}$

Regresión Lineal

$\boldsymbol{Y = \beta_0 + \beta_1 X + e}$

Supuestos:
$E(e_i) = 0; Var(e_i) = \sigma^2;$
$Cov(e_i, e_j) = 0 \; \forall \; i \neq j; e_i \sim N(0, \sigma^2)$

Estimación de parámetros

$e_i = y_i - \widehat{y_i}$

$Hallar\ \widehat{\beta_0}\ y\ \widehat{\beta_1}\ con\ SCE = \sum \widehat{e_i}^2$

$\widehat{\beta_0} = \bar{y} - \widehat{\beta_1}\bar{x} \qquad \widehat{\beta_1} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2}$

Propiedades de los estimadores (Centrados)

$\mathrm{E}(\hat{\beta}_1) = \beta_1$

$Var(\beta_1) = \frac{\sigma^2}{S_{xx}} \quad donde\ S_{XX} = \sum(x_i - \bar{x})$

Prueba asociada

$SCT = SCR + SCE$

$\sum(y_i - \bar{y})^2 = \sum(\widehat{y_i} - \bar{y})^2 + \sum(y_i - \widehat{y_i})^2$

$F_{(1;n-2)} \rightarrow \frac{SCR}{SCE/n-2}$

$R^2 = \frac{SCR}{SCT} = 1 - \frac{SCE}{SCT} \qquad 0 \leq R^2 \leq 1$

$\boldsymbol{Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_j X_j + e}$

Estimación de parámetros

$\boldsymbol{\widehat{e_i}} = y_i - \left(\widehat{\beta_0} + \widehat{\beta_1}x_i + \cdots + \widehat{\beta_j}x_{ij}\right)$

Prueba asociada

$SCT = SCR + SCE$

Recuerde:
Prueba significancia global $\rightarrow F$
Prueba significancia individual $\rightarrow t$

$t_{(gl\ error)} = \frac{\widehat{\beta_i}}{des(\widehat{\beta_i})}$

$F_{(q;n-q-1)} \rightarrow \frac{\frac{SCR}{q}}{SCE/n-q-1} = \frac{MCR}{MCE}$

Intervalo de Confianza

$IC_{(1-\alpha)}(\beta_i) = \hat{\beta} \pm t_{\left(1-\frac{\alpha}{2}, n-q-1\right)}\ d.e(\hat{\beta})$