RESUMEN PARCIAL 3 ESTADÍSTICA I

Distribuciones bivariadas Discretas:
$P(x,y) = P\left(X = x, Y = y\right)$
$P(x,y) \geq 0 \qquad \sum_x \sum_y p(x,y) = 1 \qquad SiA \subseteq R, P\left((x,y) \in A\right) = \sum_A \sum_A p(x,y)$
Distribuciones bivariadas Continuas:
$f(x,y) = f\left(X = x, Y = y\right)$
$P(x,y) \geq 0 \qquad \int_x \int_y p(x,y) = 1 \qquad SiA \subseteq R, f\left((x,y) \in A\right) = \int \int_A f(x,y)$
Distribuciones Marginales
$P_x(x) = \sum_y p(x,y) \qquad P_y(y) = \sum_x p(x,y)$ Discretas
$f_x(x) = \int_{-\infty}^{\infty} f(x,y)dy \qquad f_y(y) = \int_{-\infty}^{\infty} f(x,y)dx$ Contninuas
Distribuciones Condicionales
$P_{x|y}(x|y) = \frac{p(x,y)}{p_y(y)} \qquad P_{y|x}(y|x) = \frac{p(x,y)}{p_x(x)}$ Discretas
$f_{x|y}(x|y) = \frac{f(x,y)}{f_y(y)} \qquad f_{y|x}(y|x) = \frac{f(x,y)}{f_x(x)}$ Continuas
Valores Esperados: $h(x,y)$ función de $x, yE[h(x,y)] = \sum_x \sum_y h(x,y)p(x,y)$
Covarianza:
$Cov[X,y] = E[(X - \mu_x)(Y - \mu_y)] = \quad E[XY] - E[X]E[Y]$
Correlación:
$\rho = \rho_{x,y} = Corr[X,Y] = \frac{Cov[X,y]}{\sigma_x \sigma_y}$

TEOREMAS IMPORTANTES:

1. $Cov[aX + b, cY + d] = acCov[X,Y], \quad ac \neq 0$
2. $Corr[aX+b, cY+d] = Corr[X,Y], \quad si \quad ac > 0, \quad -Corr[X,Y] \quad si \quad ac < 0$
3. Sean $X_1, ..., X_n$ una m.a y sea $Y = \sum_{i=1}^{n} c_i X_i$, **entonces:**
$V[Y] = \sum_{i=1}^{n} c_i^2 V[X_i] + 2 \sum_{i<j} \sum c_i cj Cov[X_i, X_j]$
4. Si X, Y son independientes, **entonces** $Cov[X,Y] = 0$

Sean X, y dos v.a conjuntamente distribuidas con f.m.p.c ó f.d.p.c se dice que son independientes si:
$p(x,y) = p_x(x)p_y(y) \qquad p_{y|x}(y|x) = p_y(y) \quad p_{x|y}(x|y) = p_x(x)$ Discretas
$f(x,y) = f_x(x)f_y(y) \qquad f_{y|x}(y|x) = f_y(y) \quad f_{x|y}(x|y) = f_x(x)$ Continuas

Distribuciones Muestrales Recuerde que para que X_i para $i = 1, .., n$ sea una m.a debe cumplir que cada X_i sea independiente y cada X_i tenga la misma distribución, es decir i, i, d
$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2$
Distribución de la media muestral: Sea $X_i, .., X_n$ una m.a de una distribución con media μ y varianza σ^2, entonces:

$E[\bar{X}] = \mu \qquad V[\bar{X}] = \frac{\sigma^2}{n}$ Sí $T_o = \sum_{i=1}^{n} X_i$ entonces $E[T_o] = n\mu \quad V[T_o] = n\sigma^2$

TEOREMA

Sea $X_1, .., X_n$ una m.a de una población $N(\mu, \sigma^2)$, entonces para cualquier muestra de tamaño n $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$. Tambien $T_o \sim N(n\mu, n\sigma^2)$

TEOREMA DEL LIMITE CENTRAL

Sea $X_1, ..., X_n$ una m.a de una población con media μ y varianza $\sigma^2 < \infty$, entonces si n es suficientemente grande $n \geqslant 30$ $\hat{X} \sim N(\mu, \frac{\sigma^3}{n})$ aproximadamente. Además $T_o \sim N(n\mu, n\sigma^2)$

Estimadores Puntuales: Valores aproximados para un parámetro poblacional, queremos que represente lo mejor posible a la población.

Propiedades deseables:

1. Insesgado: se dice que un estimador $\hat{\theta}$ es insesgado para un parámetro poblacional θ Sí $E[\hat{\theta}] = \theta$. En caso contrario diremos que el estimador es sesgado y su sego está dado por: $\beta = E[\hat{\theta}] - \theta$
2. Mínima varianza, si dos estimadores son insesgados escogeremos aquel que nos permita apreciar menor varianza.
3. Eficiencia: Entre dos estimadores se prefiere el que tenga un error cuadrático, el cual se calcula como: $E.C.M[\hat{\theta}] = E[(\hat{\theta} - \theta)^2] = V[\hat{\theta}] + \beta^2$

Recuerde además, que el error estándar del estimador viene dado por: $\sqrt{V[\hat{\theta}]}$

Elaborado por Yovany Ocampo Naranjo.
Estudiante de estadística Universidad Nacional de Colombia Sede Medellín.