Fórmulas para Bioestadística

Mediana: M_e

V.Discreta: $F(x) = \frac{1}{2}$

V. Continua: $Me = e_{i-1} + \frac{\frac{1}{2}N - N_{i-1}}{n_i} a_i$

Cuantiles: $C_\alpha = e_{i-1} + \frac{\alpha N - N_{i-1}}{n_i} a_i$

Moda: M_o

$$Mo = e_{i-1} + \frac{h_i - h_{i-1}}{(h_i - h_{i-1}) + (h_i - h_{i+1})} a_i$$

Momentos

$$\mu_r = \frac{\sum_{i=1}^{k} n_i (x_i - \bar{x})^r}{n}$$

Análisis datos bidimensionales: Regresión y correlación

$$b = \frac{Cov[X,Y]}{Var[X]} = \frac{\sigma_{xy}}{\sigma_x^2} = \frac{\frac{1}{n}\sum n_i x_i y_i - \overline{x}\,\overline{y}}{\frac{1}{n}\sum n_i x_i^2 - \overline{x}^2}$$

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

$$y = a + bx$$

Teorema de la Probabilidad Compuesta

$$P(A \cap B \cap C) = P(A)\,P(B/A)\,P(C/A \cap B)$$

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

Teorema de la Probabilidad total

$$P(B) = \sum_{i=1}^{n} P(A_i)\,P(B/A_i)$$

Teorema de Bayes

$$P(A_i/B) = \frac{P(A_i)\,P(B/A_i)}{\sum_{i=1}^{n} P(A_i)\,P(B/A_i)}$$

Distribución de la media muestral $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

Varianza poblacional, σ^2, conocida

$$\overline{X} \to N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \Longrightarrow Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \to N(0,1) \qquad \left[\overline{X} - z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}},\ \overline{X} + z_{\alpha/2}\frac{\sigma_0}{\sqrt{n}}\right]$$

Varianza poblacional, σ^2, desconocida

$$T = \frac{\overline{X} - \mu}{S/\sqrt{n}} \to t_{n-1} \qquad \left[\overline{X} - t_{n-1,\alpha/2}\frac{S}{\sqrt{n}},\ \overline{X} + t_{n-1,\alpha/2}\frac{S}{\sqrt{n}}\right]$$

Varianza poblacional, σ^2, desconocida. Muestras grandes, $n > 30$

$$Z = \frac{\overline{X} - \mu}{S/\sqrt{n}} \to N(0,1) \qquad \left[\overline{X} - z_{\alpha/2}\frac{S}{\sqrt{n}},\ \overline{X} + z_{\alpha/2}\frac{S}{\sqrt{n}}\right]$$

Distribución de la varianza muestral $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$

Media poblacional, μ, desconocida

$$\chi^2 = \frac{n\hat{\sigma}^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \to \chi^2_{n-1} \qquad \left[\frac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}},\ \frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha/2}}\right]$$

Distribución del cociente de varianzas muestrales

$$\frac{S_x^2}{S_y^2} = \frac{\frac{1}{n_x - 1}\sum_{i=1}^{n_x}(X_i - \overline{X})^2}{\frac{1}{n_y - 1}\sum_{i=1}^{n_y}(Y_i - \overline{Y})^2}$$

$$F = \frac{S_x^2/\sigma_x^2}{S_y^2/\sigma_y^2} \to F_{n_x-1,n_y-1} \qquad \left[\frac{1}{F_{n_X-1,n_Y-1;\alpha/2}}\frac{S_x^2}{S_y^2},\ \frac{1}{F_{n_x-1,n_y-1;1-\alpha/2}}\frac{S_x^2}{S_y^2}\right]$$

Distribución de la diferencia de medias muestrales

$$\overline{X} - \overline{Y} = \frac{1}{n_x}\sum_{i=1}^{n_x} X_i - \frac{1}{n_y}\sum_{i=1}^{n_y} Y_i$$

Varianzas poblacionales conocidas

$$\overline{X} - \overline{Y} \to N\left(\mu_x - \mu_y, \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}\right) \qquad \left[(\overline{X} - \overline{Y}) \pm z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}\right]$$

Varianzas poblacionales desconocidas, pero iguales $(\sigma_x^2 = \sigma_y^2)$

$$S_p = \sqrt{\frac{(n_x - 1)S_x^2 + (n_y - 1)S_y^2}{n_x + n_y - 2}}$$

$$\frac{\overline{X} - \overline{Y} - (\mu_x - \mu_y)}{S_p\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \to t_{n_x+n_y-2} \qquad \left[\overline{X} - \overline{Y} \pm t_{n_X+n_Y-2,\alpha/2} \times S_p\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}\right]$$

Varianzas poblacionales desconocidas distintas o no, con n_x, $n_y > 30$

$$\frac{\overline{X} - \overline{Y} - (\mu_x - \mu_y)}{\sqrt{\frac{S_x^2}{n_x} + \frac{S_y^2}{n_y}}} \to N(0,1) \qquad \left[(\overline{X} - \overline{Y}) \pm z_{\alpha/2}\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}\right]$$

Intervalos

$$P\left[h_1(\overline{X}) \le \mu \le h_2(\overline{X})\right] = 1 - \alpha$$

Varianza y quasivarianza muestral

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 \qquad S^2 = \frac{n}{n-1}\hat{\sigma}^2$$

Distribución de la proporción muestral (Binomial n > 30)

$$X \to N\left(np, \sqrt{np(1-p)}\right)$$

Definimos el estadístico proporción muestral como: $\hat{P} = \frac{X}{n}$

$$\hat{P} \to N\left(p, \sqrt{\frac{p(1-p)}{n}}\right) \qquad \left[\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\ \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

Distribución para la diferencia de proporciones muestrales (Binomiales n > 30)

$$X \to N\left(n_x p_x, \sqrt{n_x p_x (1 - p_x)}\right) \qquad Y \to N\left(n_y p_y, \sqrt{n_y p_y (1 - p_y)}\right)$$

$$\hat{P}_x = \frac{X}{n_x} \qquad \hat{P}_y = \frac{Y}{n_y}$$

$$\hat{P}_x - \hat{P}_y \to N\left(p_x - p_y, \sqrt{\frac{p_x(1-p_x)}{n_x} + \frac{p_y(1-p_y)}{n_y}}\right)$$

$$\left[(\hat{p}_x - \hat{p}_y) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_x(1-\hat{p}_x)}{n_X} + \frac{\hat{p}_y(1-\hat{p}_y)}{n_Y}}\right]$$

Contraste para la bondad de ajuste

$$Y = \sum_{i=1}^{k}\frac{(n_i - n p_i)^2}{n p_i} \to \chi^2_{(k-1)-r}$$

Valores esperados:

$$e_{ij} = \frac{n_{i.} n_{.j}}{n}$$

Contraste para la independencia de caracteres
Contraste de homogeneidad

$$U = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(n_{ij} - e_{ij})^2}{e_{ij}} \to \chi^2_{(r-1)(s-1)}$$