Descriptive Statistics

- $\bar{x} = \frac{\sum x}{n}$ or $\bar{x} = \frac{\sum xf}{\sum f}$
- $s = \sqrt{\frac{\sum x^2 - n\bar{x}^2}{n-1}}$ or $s = \sqrt{\frac{\sum x^2 f - n\bar{x}^2}{n-1}}$
- $R_\alpha = \frac{\alpha(n+1)}{100}$

  & $P_\alpha = X_{(i)} + d\left(X_{(i+1)} - X_{(i)}\right)$

The one-sample problem

P-value approach

Hypothesis type	P – value	
Lower tail	$P(Z < z)$	$P(T_v < t)$
Upper tail	$P(Z > z)$	$P(T_v > t)$
2-tailed	$2P(Z > \|z\|)$	$2P(T_v > \|t\|)$

Test statistics

$Z_0 = \frac{(\bar{x} - \mu_0)\sqrt{n}}{\sigma}$ or $T_0 = \frac{(\bar{x} - \mu_0)\sqrt{n}}{s}$

$Z_0 = \frac{\hat{p} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$ where $\hat{p} = \frac{x}{n}$

For testing hypotheses about *σ₁ – σ₂*

Test statistic

$F_0 = \frac{s_i^2}{s_j^2}$ where $\begin{matrix} fd_1 = n_i - 1 \\ fd_2 = n_j - 1 \end{matrix}$

The two-sample problem

The i^th **paired difference** $d_i = x_{1i} - x_{2i}$ &

$T_0 = \frac{(\bar{d} - \mu_0)\sqrt{n}}{s_d}$ & $\bar{d} = \frac{\sum d}{n}$ & $s_d = \sqrt{\frac{\sum d^2 - n\bar{d}^2}{n-1}}$

Test statistics

- $Z_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$
- $Z_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$
- $T_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ & $s_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}$

$T_0 = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_1^2}{n_2}}}$ & $df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$

To test hypotheses about *π₁ - π₂*

$Z_0 = \frac{(\hat{p}_1 - \hat{p}_2) - \pi_0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

where $\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$ & $\hat{p}_1 = \frac{x_1}{n_1}$ $\hat{p}_2 = \frac{x_2}{n_2}$

Test of Independence in *rxc* table

$\chi_0^2 = \sum_{j=1}^{r}\sum_{i=1}^{c}\frac{(o_{ij} - e_{ij})^2}{e_{ij}}$ $df = (c-1)(r-1)$

Marascuillo's Test for Pair-wise Proportions

$|p_i - p_j| > \sqrt{\chi_\alpha^2}\sqrt{\left(\frac{p_i(1-p_i)}{n_i} + \frac{p_j(1-p_j)}{n_j}\right)}$

McNemar Test (Related Samples)

Test statistic $Z_0 = \frac{B - C}{\sqrt{B + C}}$

For testing hypotheses about *σ*

Test statistic $\chi_0^2 = \frac{(n-1)s^2}{\sigma_0^2}$

Regression

Sample correlation coefficient

$r = \frac{Sxy}{\sqrt{SxxSyy}}$ where $Sxx = \sum(x - \bar{x})^2$

$Syy = \sum(y - \bar{y})^2$ and $Sxy = \sum(x - \bar{x})(y - \bar{y})$

For testing ρ

Test statistic $T_0 = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ & $df = n - 2$

Estimated regression model

$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

where $\hat{\beta}_1 = \frac{Sxy}{Sxx}$ & $\hat{\beta}_0 = \bar{y} - b_1\bar{x}$

Total Sum of Squares

$SST = Syy = \sum(y - \bar{y})^2 = \sum y^2 - n\bar{y}^2$

$SSR = \hat{\beta}_1 Sxy = \frac{Sxy^2}{Sxx}$ & $SSE = SST - SSR$

Coefficient of Determination

$R^2 = \frac{SSR}{SST}$ and $R_{adj}^2 = 1 - (1 - R^2)\left(\frac{n-1}{n-k-1}\right)$

Standard Error of the model

$S_\epsilon = \sqrt{\hat{\sigma}^2} = \sqrt{MSE} = \sqrt{\frac{SSE}{n-k-1}}$

Standard Error of the Slope

$S_e(\hat{\beta}_1) = \frac{S_\epsilon}{\sqrt{Sxx}}$

For testing β_1
The test statistic & **C.I. for the slope**

$$T_0 = \frac{\hat{\beta}_1 - \beta_{10}}{S_e(\hat{\beta}_1)} \quad where \quad df = n - k - 1$$

and $\hat{\beta}_1 \pm t_{\frac{\alpha}{2},df}\, S_e(\hat{\beta}_1)$

C.I. for the **mean** of y given a particular x_p

$$\hat{y} \pm t_{\frac{\alpha}{2},df}\, S_\epsilon \sqrt{\frac{1}{n} + \frac{(x_p - \bar{x})^2}{Sxx}}$$

P.I. estimate for an **Individual value of y** given a particular x_p

$$\hat{y} \pm t_{\frac{\alpha}{2},df}\, S_\epsilon \sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{Sxx}}$$

For testing
H_0: $\beta_1 = \beta_2 = \ldots = \beta_k = 0$ H_1: at least one $\beta_i \neq 0$
Test statistic

$$F_0 = \frac{MSR}{MSE} \quad where \quad \begin{matrix} df_1 = k \\ df_2 = n - k - 1 \end{matrix}$$

Contribution of a Single Independent Variable X_j

$$SSR(X_j | all\ other\ X's) = SSR_{Full} - SSR_{(X_j)}$$

$$r^2_{Y2.1} = \frac{SSR(X_j | all\ other\ X's)}{SST_{Full} - SSR_{Full} + SSR(X_j | all\ other\ X's)}$$

The Partial F-Test Statistic

$$F_0 = \frac{SSR(X_j | all\ other\ X's)}{MSE_{Full}} \quad where \quad \begin{matrix} df_1 = 1 \\ df_2 = n - k_{full} - 1 \end{matrix}$$

For testing H_0: $\beta_{j+1} = \beta_{j+2} = \ldots\ \beta_{j+m} = 0$
against H_1: at least one $\beta_i \neq 0$

Test statistic $F_{stat} = \frac{\frac{SSR_{Full} - SSR_{Reduced}}{m}}{MSE_{Full}}$

$$where \quad \begin{matrix} df_1 = m = k_{Full} - k_{Reduced} \\ df_2 = n_{Full} - k_{full} - 1 \end{matrix}$$

Variance Inflationary Factor $VIF_j = \frac{1}{1 - R_j^2}$

$$C_p = \frac{(1 - R_k^2)(n - T)}{1 - R_T^2} - (n - 2k - 2)$$

Simple Index number formula & Unweighted aggregate price index formula (respectively)

$$I_t = \frac{y_t}{y_0}(100) \quad \& \quad I_t = \frac{\sum p_t}{\sum p_0}(100)$$

Weighted Aggregate Price Indexes

Paasche $I_t = \frac{\sum q_t p_t}{\sum q_t p_0}(100)$

Laspeyres $I_t = \frac{\sum q_0 p_t}{\sum q_0 p_0}(100)$

$$y_{adj} = \frac{y_t}{I_t}(100)$$

Single Exponential Smoothing Model

$$E_{t+1} = w y_t + (1 - w) E_t$$

Exponential Trend Model

$$y_t = \beta_0 \beta_1^{x_t} \varepsilon_t$$

Transformed Exponential Trend Model

$$\log(y_t) = \log(\beta_0) + x_t \log(\beta_1) + \log(\varepsilon_t)$$

Exponential Model for Quarterly data

$$y_t = \beta_0 \beta_1^{x_t} \beta_2^{Q_1} \beta_3^{Q_2} \beta_4^{Q_3} \varepsilon_t$$