CI for difference in proportions:

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}.$$

Hypothesis test for proportions:

Step1) Construct test statistic:

$$Z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

Step 2. Define $\mathcal{C} = (-\infty, z_{\alpha/2}) \cup (z_{\alpha/2}, \infty)$. Then reject H_0 if $Z \in \mathcal{C}$ and accept H_0 if $Z \notin \mathcal{C}$.

Since

$$0.0718 = p\text{-value} > \alpha = 0.05,$$

we do not reject H_0 at the level $\alpha = 0.05$.

Hypothesis testing for difference of proportions:

Two populations Y_1 and Y_2 are independent. Test

$$H_0 : p_1 = p_2 \qquad \text{V.S.} \qquad H_1 : p_1 \neq p_2.$$

- Under H_0, test statistics:

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}, \qquad \hat{p} = \frac{Y_1 + Y_2}{n_1 + n_2} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}.$$

18. If $E(X) = 75$ and $E(X^2) = 5949$, use Chebyshev's inequality to determine

$\mu = E(X) = 75$, $\qquad \sigma^2 = \text{Var}(X) = 5949 - 75^2 = 324.$

$\sigma = SD(X) = 18.$

a) A lower bound for $P(0 < X < 150)$.

By Chebyshev's Inequality, for $\varepsilon > 0$, $\qquad P(|X - \mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2}$

$P(0 < X < 150) = P(|X - 75| < 75) \geq 1 - \frac{324}{75^2} = \mathbf{0.9424}.$

b) A lower bound for $P(30 < X < 120)$.

$P(30 < X < 120) = P(|X - 75| < 45) \geq 1 - \frac{324}{45^2} = \mathbf{0.84}.$

c) An upper bound for $P(|X - 75| \geq 30)$.

By Chebyshev's Inequality, for $\varepsilon > 0$, $\qquad P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$

$P(|X - 75| \geq 30) \leq \frac{324}{30^2} = \mathbf{0.36}.$

Hypothesis testing for one mean:

Known σ^2:

test statistic

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

H_0	H_1	Critical Region
$\mu = \mu_0$	$\mu > \mu_0$	$z \geq z_\alpha$
$\mu = \mu_0$	$\mu < \mu_0$	$z \leq -z_\alpha$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert z\rvert \geq z_{\alpha/2}$

Unknown σ^2:

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \qquad S = \sqrt{S^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

H_0	H_1	Critical Region
$\mu = \mu_0$	$\mu > \mu_0$	$t \geq t_\alpha(n-1)$
$\mu = \mu_0$	$\mu < \mu_0$	$t \leq -t_\alpha(n-1)$
$\mu = \mu_0$	$\mu \neq \mu_0$	$\lvert t\rvert \geq t_{\alpha/2}(n-1)$

Hypothesis testing for variance:

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \qquad S^2 = s^2 = (n-1)^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Confidence Interval for two means (known variances):

$$(\bar{X} - \bar{Y}) \pm z_{\alpha/2}\sigma_\delta, \quad \sigma_\delta = \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$$

Confidence Interval for two means (Unknown variances):

$$(\bar{X}-\bar{Y}) \pm z_{\alpha/2} S_{\delta}, \quad S_{\delta} = \sqrt{S_1^2/n_1 + S_2^2/n_2}$$

Confidence Interval for two means (unknown equal variances):

$$t_0 = t_{\alpha/2}(n_1+n_2-2) \qquad S_{\text{pooled}}^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$$

$$(\bar{X}-\bar{Y}) \pm t_0 S_{\text{pooled}} \sqrt{n_1^{-1} + n_2^{-1}}$$

Hypothesis test for two means (unknown equal variances):

$$T = \frac{(\bar{X}-\bar{Y}) - \delta_0}{S_{\text{pooled}} \sqrt{n_1^{-1} + n_2^{-1}}}$$

Decision rule: reject H_0 if $|T| \geq t_0$

Hypothesis test for two means (unknown unequal variances):

$$T \approx \frac{(\bar{X}-\bar{Y}) - \delta}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$$

Paired comparison:

$$\bar{\delta} = \frac{1}{n}\sum_{i=1}^{n} \delta_i \quad \text{and} \quad S_{\delta}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\delta_i - \bar{\delta})^2$$

$100(1-\alpha)\%$ **Confidence interval**:

$$\bar{\delta} \pm t_{\alpha/2}(n-1)\frac{S_{\delta}}{\sqrt{n}}.$$

Hypothesis testing at significance level α:

$$H_0 : \delta = \delta_0 \qquad \text{V.S.} \qquad H_1 : \delta \neq \delta_0.$$

Test statistic:

$$\Delta = \frac{\bar{\delta} - \delta_0}{S_{\delta}/\sqrt{n}}.$$

Decision rule: reject H_0 if $|\Delta| \geq t_{\alpha/2}(n-1)$.

Common Series for Probability: (Taylor Series) $f(x) = \sum_{n=0}^{\infty} \frac{(x-a)^n}{n!} f^{(n)}(a)$

(Geometric Series) $\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$ $\qquad \sum_{k=m}^{\infty} x^k = \frac{x^m}{1-x}$

(Exponential Series)- $\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$

e) Construct a 95% (two-sided) confidence interval for the overall standard deviation.

Confidence Interval for σ^2: $\left(\frac{(n-1)\cdot s^2}{\chi^2_{\alpha/2}}, \frac{(n-1)\cdot s^2}{\chi^2_{1-\alpha/2}} \right)$.

$\alpha = 0.05.$ $\qquad \alpha/2 = 0.025.$ $\qquad 1-\alpha/2 = 0.975.$

number of degrees of freedom $= n - 1 = 9 - 1 = 8.$

$\chi^2_{\alpha/2} = 17.54.$ $\qquad \chi^2_{1-\alpha/2} = 2.180.$

$\left(\frac{(9-1)\cdot 0.3025}{17.54}, \frac{(9-1)\cdot 0.3025}{2.180} \right)$ $\qquad$ (0.13797 ; 1.11009)

Confidence Interval for σ: $\left(\sqrt{0.13797}, \sqrt{1.11009}\right) =$ **(0.3714 ; 1.0536)**

f) Construct a 90% one-sided confidence interval for σ that provides an upper bound for σ.

$\left(0, \sqrt{\frac{(n-1)\cdot s^2}{\chi^2_{1-\alpha}}} \right)$ $\qquad \chi^2_{1-\alpha} = \chi^2_{0.90} = 3.490.$

$\left(0, \sqrt{\frac{(9-1)\cdot 0.3025}{3.490}} \right)$ $\qquad$ **(0 , 0.8327)**

g) Construct a 95% one-sided confidence interval for σ that provides a lower bound for σ.

$\left(\sqrt{\frac{(n-1)\cdot s^2}{\chi^2_{\alpha}}}, \infty \right)$ $\qquad \chi^2_{\alpha} = \chi^2_{0.05} = 15.51.$

$\left(\sqrt{\frac{(9-1)\cdot 0.3025}{15.51}}, \infty \right)$ $\qquad$ **(0.395 , ∞)**

Covariance and Correlation Coefficient:

$$E[u(X,Y)] = E[(X-\mu_X)(Y-\mu_Y)] = \sigma_{XY} = \text{Cov}(X,Y)$$

σ_X and σ_Y are positive, then

$$\rho = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

$$\mu_X = E(X) = \sum_x \sum_y x f(x,y) = \sum_x x\left[\sum_y f(x,y)\right] = \sum_x x f_X(x).$$

$$\text{Cov}(X,Y) = E(XY) - \mu_X\mu_Y - \mu_Y\mu_X + \mu_X\mu_Y = E(XY) - \mu_X\mu_Y.$$