Jane Downer
April 24, 2021

Homework 5

Problem 1 A large-sample α-level test of hypothesis for

$$\begin{cases} H_0: & \theta = \theta_0 \\ & vs \\ H_a: & \theta > \theta_0 \end{cases}$$

rejects the null hypothesis if

$$\frac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}} > Z_\alpha$$

Show that this is equivalent to the following rejection criterion: Reject the null if θ_0 is less than the large-sample $100(1-\alpha)\%$ lower confidence bound for θ.

Show:

$\dfrac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}} > z_\alpha$ is equivalent to $\theta_0 <$ lower bound

Find lower bound for upper-tail $100(1-\alpha)\%$ confidence interval:

$$\text{confidence interval} = [\underbrace{\hat{\theta} - z_\alpha \sigma_{\hat{\theta}}}_{\text{lower bound}}, \infty)$$

Show $\theta_0 <$ lower bound is equivalent to original expression:

$$\theta_0 < \hat{\theta} - z_\alpha \sigma_{\hat{\theta}}$$

$$\shortparallel$$

$$\hat{\theta} - z_\alpha \sigma_{\hat{\theta}} > \theta_0$$

$$\shortparallel$$

$$\hat{\theta} - \theta_0 > z_\alpha \sigma_{\hat{\theta}}$$

$$\shortparallel$$

$$\frac{\hat{\theta} - \theta_0}{\sigma_{\hat{\theta}}} > z_\alpha \quad \checkmark\checkmark$$

Problem 2 Two groups of children from an elementary school were taught reading by using different methods. A reading test on 49 children from each group yielded the results $\bar{y}_1 = 75, \bar{y}_2 = 71, s_1 = 9,$ and $s_2 = 10$.

1) If you wish to conclude whether evidence indicates a difference between the two population means, how would you construct the hypotheses?
2) Use the p-value approach to perform hypothesis testing. What would you conclude if you desired an α value of 0.05?
3) Method 1 yielded better average score than method 2. If you wish to find out if method 1 is better, what would be the hypotheses, and what would you conclude at level 0.05?

1.

$$\boxed{\begin{cases} H_0: \mu_1 - \mu_2 = 0 \\ H_a: \mu_1 - \mu_2 \neq 0 \end{cases}}$$

2. samples sufficiently large to use z-test

$$z = \left[(\bar{y}_1 - \bar{y}_2) - 0\right] \Big/ \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= \left[(75 - 71) - 0\right] \Big/ \sqrt{\frac{9^2}{49} + \frac{10^2}{49}}$$

$$= 7 \times 4 / \sqrt{181}$$

$$= 2.081$$

$$RR: \{p : p < \alpha\}$$

where $p = P(-z_{\alpha/2} < -2.081) + P(z_{\alpha/2} > 2.081)$

$= 0.0188 + 0.0188$

$= 0.0376$

$0.0376 < 0.05$

$\therefore$ reject null hypothesis that the two population means are equal

3.

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_a: \mu_1 > \mu_2 \end{cases}$$

d_0 = difference in population means under the null (0)

$$\text{z-statistic} = \left[(\bar{y}_1 - \bar{y}_2) - d_0\right] \Big/ \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$$= \left[(75 - 71) - 0\right] \Big/ \sqrt{\frac{9^2}{49} + \frac{10^2}{49}}$$

$$= 4 \times 7 \big/ \sqrt{181}$$

$$= 2.081$$

$$RR: \{p : p < \alpha\}$$

where $p = P(Z_\alpha > 2.081)$

$= 0.0188$

$0.0188 < 0.05$

$\therefore$ reject null hypothesis that the two population means are equal

10.86 Aptitude tests should produce scores with a large amount of variation so that an administrator can distinguish between persons with low aptitude and persons with high aptitude. The standard test used by a certain industry has been producing scores with a standard deviation of 10 points. A new test is given to 20 prospective employees and produces a sample standard deviation of 12 points. Are scores from the new test significantly more variable than scores from the standard? Use $\alpha = .01$.

$$\begin{cases} H_0: \sigma = 10 \\ H_a: \sigma > 10 \end{cases}$$

$n = 20$

$S = 12$

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{19 \cdot 12^2}{10^2} = 27.36$$

$$RR: \{\chi^2 : \chi^2 > \chi^2_{\alpha,v}\}$$

v = degrees of freedom = $20-1 = 19$

$\alpha = 0.01$

From table: $\chi^2_{0.01,\,19} = 36.1908$

Rejection criteria not satisfied — 27.36 is not larger than 36.1908

$$\boxed{\therefore \text{ No, the new scores are not significantly more variable.}}$$

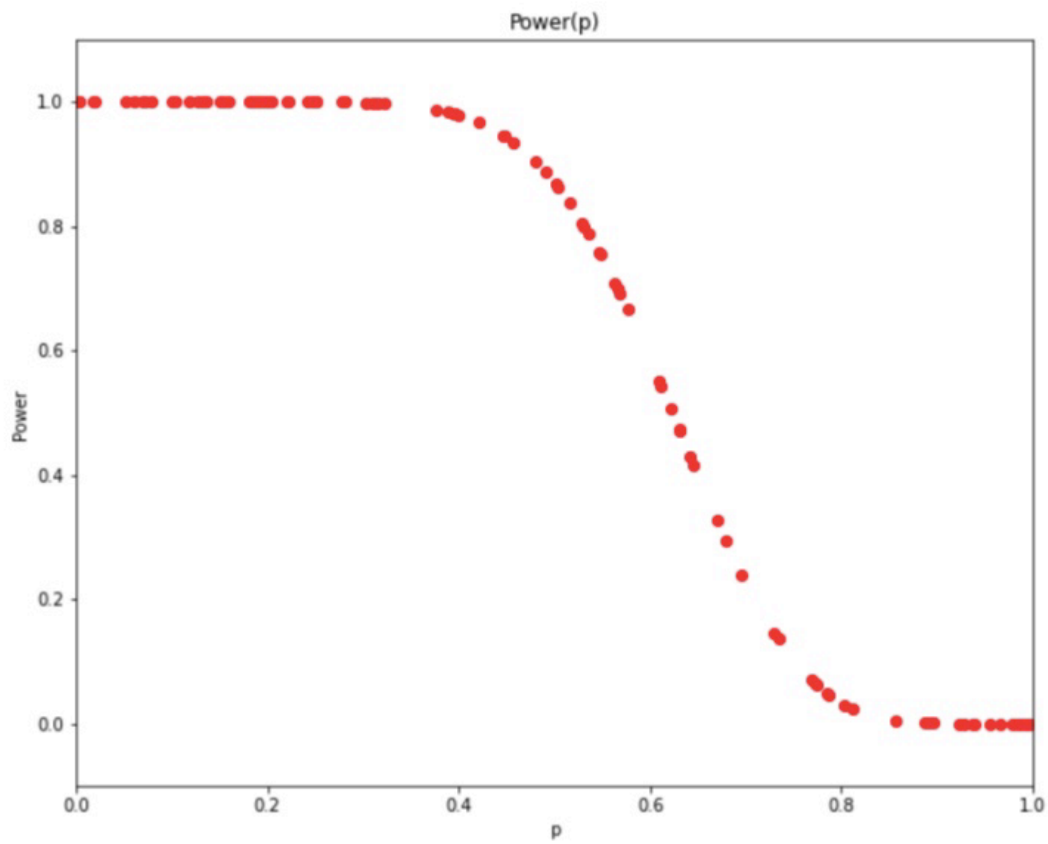
Problem 4 For the application described in Exercise 10.2 (on page 494), derive the power function, $power(p)$, as a function of p, then sketch a graph of the power function as p varies. Note that $0 \leq p \leq 1$. Use 100 data points between 0 and 1 to plot the graph by using any software of your choice.

$$\begin{cases} H_0: p' = p \\ H_a: p' < p \end{cases} \qquad n = 20$$

RR: $\{y : y \leq 12\}$

$\beta = P(\text{fail to reject } H_0 \mid H_a \text{ is true})$

$$\begin{aligned} \text{power} &= 1 - \beta \\ &= P(\text{reject } H_0 \mid H_a) \\ &= P(y \leq 12 \mid p' < p) \\ &= \sum_{y=0}^{12} \binom{20}{y} p^{y} \cdot (1-p)^{20-y} \end{aligned}$$



10.96 Suppose Y is a random sample of size 1 from a population with density function

$$f(y \mid \theta) = \begin{cases} \theta y^{\theta-1}, & 0 \le y \le 1, \\ 0, & \text{elsewhere,} \end{cases}$$

where $\theta > 0$.

a Sketch the power function of the test with rejection region: $Y > .5$.

b Based on the single observation Y, find a uniformly most powerful test of size α for testing $H_0 : \theta = 1$ versus $H_a : \theta > 1$.

(a) $\begin{cases} H_0: \theta = \theta_0 \\ H_a: \theta = \theta_a \end{cases}$

$$\text{Power} = 1 - \beta$$

$$\beta = P(\text{fail to reject } H_0 \mid H_a \text{ is true})$$

$$= P(y \le 0.5 \mid \theta = \theta_a)$$

$$= \int_0^{0.5} f(y|\theta)\,dy \quad \text{where } \theta = \theta_a$$

$$= \int_0^{0.5} \theta_a y^{\theta_a - 1}\,dy$$

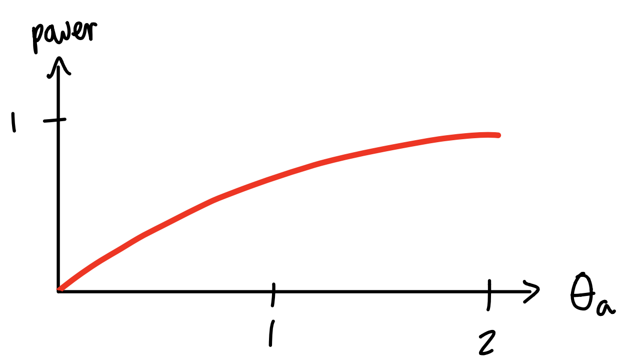
$$= \frac{1}{(\theta_a - 1) + 1} \cdot \theta_a \cdot y^{(\theta_a - 1)+1} \Big|_0^{0.5}$$

$$= y^{\theta_a} \Big|_0^{0.5}$$

$$= 0.5^{\theta_a} - 0^{\theta_a}$$

$$= 0.5^{\theta_a}$$

$$\hookrightarrow \text{power}(\theta_a) = 1 - 0.5^{\theta_a}$$



(b)

$$\begin{cases} H_0: \theta = 1 \\ H_a: \theta > 1 \end{cases}$$ single observation

$$f_Y(y|\theta) = \theta \cdot y^{\theta - 1}$$

UMP test $\rightarrow$ Neyman Pearson Lemma

$$\frac{L(\theta_a)}{L(\theta_0)} = \frac{f_Y(y|\theta_a)}{f_Y(y|\theta_0)}$$

$$= \frac{\theta_a \cdot y^{\theta_a - 1}}{\theta_0 \cdot y^{\theta_0 - 1}}$$

$$= \frac{\theta_a}{\theta_0} \cdot y^{(\theta_a - 1) - (\theta_0 - 1)}$$

$$= \theta_a \cdot y^{\theta_a - 1}$$

k = some cutoff value

$$RR = \left\{ \frac{L(\theta_a)}{L(\theta_0)} : \frac{L(\theta_a)}{L(\theta_0)} > k \right\}$$

$$= \left\{ \theta_a \cdot y^{\theta_a - 1} : \theta_a \cdot y^{\theta_a - 1} > k \right\}$$

$$= \left\{ y : y > c \right\} \longrightarrow \text{where } c = \theta_a^{-1} \cdot k^{(\theta_a - 1)^{-1}}$$

$$\alpha = P(\text{reject } H_0 | H_0 \text{ is true})$$

$$= P(y > c | \theta = 1)$$

$$= \int_c^1 \theta_0 \cdot y^{\theta_0 - 1} dy$$

$$= \int_c^1 1 \cdot dy$$

$$= y \Big|_c^1$$

$$= 1 - c \longrightarrow \alpha = 1 - c, \text{ so } c = 1 - \alpha$$

$$\boxed{\therefore RR = \{ y : y > 1 - \alpha \}}$$

10.108 Suppose that $X_1, X_2, \ldots, X_{n_1}, Y_1, Y_2, \ldots, Y_{n_2}$, and $W_1, W_2, \ldots, W_{n_3}$ are independent random samples from normal distributions with respective unknown means μ_1, μ_2, and μ_3 and variances σ_1^2, σ_2^2, and σ_3^2.

a Find the likelihood ratio test for $H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$ against the alternative of at least one inequality.

b Find an approximate critical region for the test in part (a) if n_1, n_2, and n_3 are large and $\alpha = .05$.

(a) $\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2 \\ H_a: \text{at least one inequality} \end{cases}$

$$\Theta = (\mu_1, \mu_2, \mu_3, \sigma_1^2, \sigma_2^2, \sigma_3^2)$$

$$\Omega_0 = \{(\mu_1, \mu_2, \mu_3, \sigma_1^2, \sigma_2^2, \sigma_3^2): -\infty < \mu_1, \mu_2, \mu_3 < \infty \;\&\; \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2 \;\&\; \sigma^2 > 0\}$$

$$\Omega_a = \{(\mu_1, \mu_2, \mu_3, \sigma_1^2, \sigma_2^2, \sigma_3^2): -\infty < \mu_1, \mu_2, \mu_3 < \infty \;\&\; \text{unequal variance} \;\&\; \sigma_1^2, \sigma_2^2, \sigma_3^2 > 0\}$$

$$\Omega = \{(\mu_1, \mu_2, \mu_3, \sigma_1^2, \sigma_2^2, \sigma_3^2): -\infty < \mu_1, \mu_2, \mu_3 < \infty \;\&\; \sigma_1^2, \sigma_2^2, \sigma_3^2 > 0\}$$

Get $L(\hat{\Omega}_0)$ and $L(\hat{\Omega})$

$$L(\Theta) = \prod_{j=1}^{3} \prod_{i=1}^{n_j} \left[\frac{1}{\sigma_j \sqrt{2\pi}} \cdot \exp\left(\frac{-1}{2\sigma_j^2} (x_{ij} - \mu_j)^2 \right) \right]$$

$$= \left[\prod_{j=1}^{3} \prod_{i=1}^{n_j} \frac{1}{\sigma_j \sqrt{2\pi}} \right] \cdot \exp\left(\sum_{j=1}^{3} \sum_{i=1}^{n_j} \frac{-1}{2\sigma_j^2} (x_{ij} - \mu_j)^2 \right)$$

$L(\Omega)$:

$$\text{MLE} \begin{cases} \hat{\mu}_j = \overline{X}_j \\ \hat{\sigma}_j = \frac{1}{n_j - 1} \sum_{i=1}^{n_j} (x_{ij} - \overline{x}_j)^2 \end{cases}$$

$$L(\Omega) = \left[\prod_{j=1}^{3} \prod_{i=1}^{n_j} \frac{1}{\hat{\sigma}_j \sqrt{2\pi}} \right] \cdot \exp\left(\sum_{j=1}^{3} \sum_{i=1}^{n_j} \frac{-1}{2\hat{\sigma}_j^2} (x_{ij} - \overline{x}_j)^2 \right)$$

$$= \sum_{j=1}^{3} \sum_{i=1}^{n_j} \frac{-1}{2\hat{\sigma}_j^2} (x_{ij} - \overline{x}_j)^2$$

$$= \sum_{j=1}^{3} \sum_{i=1}^{n_j} \frac{-(n_j - 1)(x_{ij} - \overline{x}_j)^2}{2 \sum_{i=1}^{n_j} (x_{ij} - \overline{x}_j)^2}$$

$$= -\frac{1}{2} \sum_{j=1}^{3} (n_j - 1)$$

$$= -\left[\sum_{j=1}^{3} n_j - 3 \right] \div 2$$

$$= \left[\prod_{j=1}^{3}\prod_{i=1}^{n_j}\frac{1}{\hat{\sigma}_j\sqrt{2\pi}}\right]\cdot\exp\left(-\frac{n-3}{2}\right) \longleftarrow n = \sum_{j=1}^{3} n_j$$

$L(\Omega_0)$:

$$\hat{\sigma}^2 = \frac{1}{n-3}\sum_{j=1}^{3}\sum_{i=1}^{n_j}(x_{ij}-\bar{x}_j)^2 \quad \leftarrow \hat{\mu}_j$$

d.f. = 3

$$= \frac{1}{n-3}\sum_{j=1}^{3}\left[\cancel{\sum_{i=1}^{n_j}(x_{ij}-\bar{x}_j)^2}\times(n_j-1)\times\frac{\hat{\sigma}_j^2}{\cancel{\sum_{i=1}^{n_j}(x_{ij}-\bar{x}_j)^2}}\right]$$

$$= \frac{1}{n-3}\sum_{j=1}^{3}(n_j-1)\hat{\sigma}_j^2$$

$$L(\Omega_0) = \left[\prod_{j=1}^{3}\prod_{i=1}^{n_j}\frac{1}{\hat{\sigma}_j\sqrt{2\pi}}\right]\cdot\exp\left(-\frac{n-3}{2}\right) \longrightarrow \text{where each } \hat{\sigma}_j^2 = \hat{\sigma}^2$$

$$= \left[\prod_{i=1}^{n_j}\frac{1}{\hat{\sigma}\sqrt{2\pi}}\right]^3\exp\left(-\frac{n-3}{2}\right)$$

$$\lambda = \frac{L(\Omega_0)}{L(\Omega)}$$

$$= \frac{\left[\prod_{j=1}^{3}\prod_{i=1}^{n_j}\frac{1}{\hat{\sigma}_j\sqrt{2\pi}}\right]\cdot\exp\left(-\frac{n-3}{2}\right)}{\left[\prod_{i=1}^{n_j}\frac{1}{\hat{\sigma}\sqrt{2\pi}}\right]^3\exp\left(-\frac{n-3}{2}\right)}$$

$$= \left[\prod_{j=1}^{3}\prod_{i=1}^{n_j}\frac{1}{\hat{\sigma}_j\sqrt{2\pi}}\right] \div \left[\prod_{i=1}^{n_j}\frac{1}{\hat{\sigma}\sqrt{2\pi}}\right]^3$$

(b) $RR = \{\lambda : -2\ln(\lambda) > \chi^2_{\alpha, r_0 - r}\}$

$= \{\lambda : -2\ln(\lambda) > \chi^2_{0.05, 2}\}$

$= \boxed{\{\lambda : -2\ln(\lambda) > 5.99147\}}$