MSA 8190 Statistical Foundations

Fall 2016

Sample Problems on Hypothesis Testing

Problem 1
The yield of a chemical process is being studied. From previous experience yield is known to be normally distributed and $\sigma = 3$. The past five days of plant operation have resulted in the following percent yields: 91.6, 88.75, 90.8, 89.95, and 91.3. Find a 95% two-sided confidence interval on the true mean yield.

Solution:

$$\bar{x} = \frac{91.6 + 88.75 + 90.8 + 89.95 + 91.3}{5} = 90.48$$

$\alpha = 1 - 0.95 = 0.05$
$\mathrm{z}_{\alpha/2} = \text{qnorm(0.05/2, lower.tail = FALSE)} = 1.959964$

$$\left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right] = \left[90.48 - 1.959964\frac{3}{\sqrt{5}}, 90.48 + 1.959964\frac{3}{\sqrt{5}}\right]$$

$\Rightarrow [87.85043, 93.10957]$

Problem 2
The heat evolved in calories per gram of a cement mixture is approximately normally distributed. The mean is thought to be 100 and the standard deviation is 2. We wish to test $H_0: \ \mu = 100$ versus $H_1: \ \mu \neq 100$ with a sample of $n = 9$ specimens.

(a) If acceptance region is defined as $98.5 \leq \bar{x} \leq 101.5$, find the type I error probability α.

(b) Find β for the case where the true mean heat evolved is 103.

(c) Find β for the case where the true mean heat evolved is 105. This value of β is smaller than the one found in part (b) above. Why?

Solution:

(a) $$z_1 = \frac{L - \mu_0}{\sigma/\sqrt{n}} = \frac{98.5 - 100}{2/\sqrt{9}} = -2.25$$

$$z_2 = \frac{U - \mu_0}{\sigma/\sqrt{n}} = \frac{101.5 - 100}{2/\sqrt{9}} = 2.25$$

$\alpha = 1 - \mathbb{P}(98.5 \le \bar{x} \le 101.5|\mu = 100) = \mathbb{P}(Z \le -2.25) + \mathbb{P}(Z > 101.5)$
$\Rightarrow \alpha = pnorm(-2.25) + pnorm(2.25, lower.tail = FALSE) = 0.02444895$

(b) $z_1 = \dfrac{L - \mu_1}{\sigma/\sqrt{n}} = \dfrac{101.5 - 103}{2/\sqrt{9}} = -2.25$
$z_2 = \dfrac{U - \mu_1}{\sigma/\sqrt{n}} = \dfrac{98.5 - 103}{2/\sqrt{9}} = -6.75$
$\beta = \mathbb{P}(98.5 \le \bar{x} \le 101.5|\mu = 103) = \mathbb{P}(-6.75 \le Z \le -2.25) = \mathbb{P}(Z \le -2.25) - \mathbb{P}(Z \le -6.75) = pnorm(-2.25) - pnorm(-6.75) = 0.01222447$

(c) $z_1 = \dfrac{L - \mu_2}{\sigma/\sqrt{n}} = \dfrac{101.5 - 105}{2/\sqrt{9}} = -5.25$
$z_2 = \dfrac{U - \mu_2}{\sigma/\sqrt{n}} = \dfrac{98.5 - 105}{2/\sqrt{9}} = -9.75$
$\beta = \mathbb{P}(98.5 \le \bar{x} \le 101.5|\mu = 105) = \mathbb{P}(-9.75 \le Z \le -5.25) = \mathbb{P}(Z \le -5.25) - \mathbb{P}(Z \le -9.75) = pnorm(-5.25) - pnorm(-9.75) = 7.604961e - 08$
The value of β is getting smaller because the true mean is getting farther away from the sample mean, so it becomes easier to detect that true mean is different than $\mu_0 = 100$.

Problem 3
A bearing used in an automotive application is suppose to have a nominal inside diameter of 1.5 inches. A random sample of 25 bearings is selected and the average inside diameter of these bearings is 1.4975 inches. Bearing diameter is known to be normally distributed with standard deviation $\sigma = 0.01$ inch.

(a) Test the hypotheses $H_0: \ \mu = 1.5$ versus $H_1: \ \mu \ne 1.5$ using $\alpha = 0.01$

(b) Compute the power of the test if the true mean diameter is 1.495 inches.

(c) What sample size would be required to detect a true mean diameter as low as 1.495 inches if we wanted the power of the test to be at least 0.9?

(d) Explain how the question in part (a) could be answered by constructing a two-sided confidence interval on the mean diameter.

Solution:

(a) $z_0 = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \dfrac{1.4975 - 1.5}{0.01/\sqrt{25}} = -1.25$
$z_{\alpha/2} = qnorm(0.01/2, lower.tail = FALSE) = 2.575829$
Since the $|z_0| < z_{\alpha/2}$ we can't reject H_0. In other words, we accept $\mu = 1.5$

(b) $d = \dfrac{\mu - \mu_0}{\sigma} = \dfrac{1.495 - 1.5}{0.01} = -0.5$
$pwr.norm.test(-0.5, n = 25, sig.level = 0.01)$
$\Rightarrow Power = 0.4697776$

(c) $pwr.norm.test(-0.5, sig.level = 0.01, power = 0.9)$
$\Rightarrow n = 59.51753 \approx 60$

(d) $99\%\ CI : \left[1.4975 - 2.575829\frac{0.01}{\sqrt{25}}, 1.4975 + 2.575829\frac{0.01}{\sqrt{25}}\right]$
$\Rightarrow [1.492348, 1.502652]$
Since $\mu_0 = 1.5$ is inside the CI, we can't reject H_0 and we accept $\mu = 1.5$.

Problem 4

If $X_1, X_2, \cdots, X_n \sim NID(\mu_1, \sigma^2)$, $Y_1, Y_2, \cdots, Y_m \sim NID(\mu_2, \sigma^2)$, and X's and Y's are independent, find the distribution of the following statistics. (If the distribution is normal, specify the mean and variance. If it has chi-square, t, and/or F distribution, specify the degrees of freedom.)

(a) $\bar{X} - \bar{Y}$

(b) $\frac{\sum_{i=1}^{n}(X_i-\mu_1)-\sum_{i=1}^{m}(Y_i-\mu_2)}{\sigma}$

(c) $\frac{\sum_{i=1}^{n}(X_i-\mu_1)^2}{\sigma^2}$

(d) $\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2}$

(e) $\frac{n(\bar{X}-\mu_1)^2}{\sigma^2}$

(f) $\frac{S_x^2}{S_y^2}$

(g) $\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2+\sum_{i=1}^{m}(Y_i-\bar{Y})^2}{\sigma^2}$

(h) $\frac{\sum_{i=1}^{n}(X_i-\mu_1)^2+\sum_{i=1}^{m}(Y_i-\mu_2)^2}{\sigma^2}$

(i) $\frac{\bar{X}-\mu_1}{S_x/\sqrt{n}}$

Solution:

(a) $\bar{X} \sim N(\mu_1, \frac{\sigma^2}{n})$, $\bar{Y} \sim N(\mu_2, \frac{\sigma^2}{m}) \Rightarrow\ \bar{X} - \bar{Y} \sim N(\mu_1 - \mu_2, \frac{\sigma^2}{n} + \frac{\sigma^2}{m})$

(b) $\frac{\sum_{i=1}^{n}(X_i-\mu_1)-\sum_{i=1}^{m}(Y_i-\mu_2)}{\sigma} = \frac{\sum_{i=1}^{n}(X_i-\mu_1)}{\sigma} - \frac{\sum_{i=1}^{m}(Y_i-\mu_2)}{\sigma}$

$\frac{\sum_{i=1}^{n}(X_i-\mu_1)}{\sigma} \sim N(0,n)$, $\frac{\sum_{i=1}^{m}(Y_i-\mu_2)}{\sigma} \sim N(0,m)$

$\Rightarrow \frac{\sum_{i=1}^{n}(X_i-\mu_1)-\sum_{i=1}^{m}(Y_i-\mu_2)}{\sigma} \sim N(0,n+m)$

(c) $\frac{(X_i-\mu_1)^2}{\sigma^2} \sim \chi^2_{(1)}$, $i \in \{1,2,\cdots,n\}$

$\Rightarrow \frac{\sum_{i=1}^{n}(X_i-\mu_1)^2}{\sigma^2} \sim \chi^2_{(n)}$

(d) $\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2} \sim \chi^2_{(n-1)}$

(e) $\frac{\bar{X}-\mu_1}{\sigma/\sqrt{n}} \sim N(0,1)$

$\Rightarrow \left(\frac{\bar{X}-\mu_1}{\sigma/\sqrt{n}}\right)^2 = \frac{n(\bar{X}-\mu_1)^2}{\sigma^2} \sim \chi^2_{(1)}$

(f) $\frac{(n-1)S_x^2}{\sigma^2} \sim \chi^2_{(n-1)}$, $\frac{(m-1)S_y^2}{\sigma^2} \sim \chi^2_{(m-1)}$

$\Rightarrow \frac{S_x^2}{S_y^2} \sim \mathcal{F}(n-1,m-1)$

(g) $\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\sigma^2} \sim \chi^2_{(n-1)}$

$\frac{\sum_{i=1}^{m}(Y_i-\bar{Y})^2}{\sigma^2} \sim \chi^2_{(m-1)}$

$\Rightarrow \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2+\sum_{i=1}^{m}(Y_i-\bar{Y})^2}{\sigma^2} \sim \chi^2_{(n+m-2)}$

(h) $\frac{\sum_{i=1}^{n}(X_i-\mu_1)^2}{\sigma^2} \sim \chi^2_{(n)}$

$\frac{\sum_{i=1}^{m}(Y_i-\mu_2)^2}{\sigma^2} \sim \chi^2_{(m)}$

$\Rightarrow \frac{\sum_{i=1}^{n}(X_i-\mu_1)^2+\sum_{i=1}^{m}(Y_i-\mu_2)^2}{\sigma^2} \sim \chi^2_{(n+m)}$

(i) $\frac{\bar{X}-\mu_1}{\sigma/\sqrt{n}} \sim N(0,1)$

$\frac{(n-1)S_x^2}{\sigma^2} \sim \chi^2_{(n-1)}$

$\Rightarrow \frac{(\bar{X}-\mu_1)\sqrt{n}/\sigma}{\sqrt{S_x^2/\sigma^2}} = \frac{\bar{X}-\mu_1}{S_x/\sqrt{n}} \sim t(n-1)$