Florida Atlantic University  Instructor: Hongwei Long
STA4443 Midterm Exam I
February 15, 2010

No books, notes or other aids, except calculators and the provided formula sheet, are allowed.

1. Let $X$ equal the weight in grams of a miniature candy bar. Assume that $X$ has a normal distribution $N(24.43, 2.20)$.

(a) Find $P(X < 24.82)$.

(b) Assume that 30 candy bars are selected at random and weighted. Let $Y$ be the number of these candy bars that weigh less than 24.82 grams. Approximate $P(4 \leq Y \leq 8)$.

(c) Let $\bar{X}$ be the sample mean of the 30 candy bars selected and weighted. Find $P(24.17 \leq \bar{X} \leq 24.82)$.

Solution. (a) By a standard transformation, we find

$$P(X < 24.82) = P\left(\frac{X - 24.43}{\sqrt{2.2}} < \frac{24.82 - 24.43}{\sqrt{2.2}}\right) = P(Z < 0.263) = 0.6026.$$ 

(b) It is easy to see that $Y \sim b(30, 0.6026)$ with mean $\mu_Y = np = 18.037$ and variance $\sigma_Y^2 = np(1 - p) = 7.184$. So, by using the approximation formula based on CLT, we have

$$P(4 \leq Y \leq 8) = P(3.5 \leq Y \leq 8.5) = P\left(\frac{3.5 - 18.037}{\sqrt{7.184}} \leq \frac{Y - 18.037}{\sqrt{7.184}} \leq \frac{8.5 - 18.037}{\sqrt{7.184}}\right) \approx P(-5.424 \leq Z \leq -3.558) \approx 0.$$ 

(c) Note that $\bar{X}$ has a normal distribution $N(24.43, 2.2/30)$, i.e. $N(24.43, 0.0733)$. We have

$$P(24.17 \leq \bar{X} \leq 24.82) = P\left(\frac{24.17 - 24.43}{\sqrt{0.0733}} \leq \frac{\bar{X} - 24.43}{\sqrt{0.0733}} \leq \frac{24.82 - 24.43}{\sqrt{0.0733}}\right) = P(-0.96 \leq Z \leq 1.44) = P(Z \leq 1.44) - P(Z \leq -0.96) = 0.9251 - 0.1685 = 0.7566.$$ 

2. Let $X_1, X_2, \cdots, X_n$ be a random sample from exponential distribution with p.d.f.

$$f(x; \theta) = \frac{1}{\theta} e^{-x/\theta}, \ 0 < x < \infty, \ 0 < \theta < \infty.$$ 

(a) Find the maximum likelihood estimator $\hat{\theta}$.

(b) Is $\hat{\theta}$ an unbiased estimator of $\theta$? (Justify your answer).

Solution. (a) The likelihood function of the random sample is given by

$$L(\theta) = f(x_1; \theta)f(x_2; \theta)\cdots f(x_n; \theta)$$

$$= \frac{1}{\theta} e^{-x_1/\theta} \frac{1}{\theta} e^{-x_2/\theta} \cdots \frac{1}{\theta} e^{-x_n/\theta}$$

$$= \left(\frac{1}{\theta}\right)^n e^{-\sum_{i=1}^{n} x_i/\theta}.$$
We have
\[ \ln L(\theta) = -n \ln \theta - \sum_{i=1}^{n} x_i / \theta, \]

\[ \frac{d \ln L(\theta)}{d\theta} = -n/\theta + \sum_{i=1}^{n} x_i / \theta^2 = 0 \]

\[ \theta = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}. \]

So, the maximum likelihood estimator of \( \theta \) is
\[ \hat{\theta} = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i. \]

(b) Note that each \( X_i \) \((i = 1, 2, \cdots, n)\) has exponential distribution with parameter \( \theta \). Then,

\[ E(X_i) = \theta. \]

Hence,
\[ E(\hat{\theta}) = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \frac{1}{n} (n \cdot \theta) = \theta, \]

which implies that \( \hat{\theta} \) is an unbiased estimator of \( \theta \).

3. Assume that the yield per acre for a particular variety of soybeans is \( N(\mu, \sigma^2) \) with \( \sigma^2 \) unknown. For a random sample of \( n = 5 \) plots, the yields in bushels per acre were 37.4, 48.8, 46.9, 55.0 and 44.0.

(a) Give point estimates for \( \mu \) and \( \sigma \), respectively.

(b) Find a 90% confidence interval for \( \mu \).

**Solution.** (a) We use \( \bar{x} \) as a point estimate for \( \mu \). Thus, we have \( \bar{x} = (37.4 + 48.8 + 46.9 + 55.0 + 44.0)/5 = 46.42 \). We use \( s \) as a point estimate of \( \sigma \). Note that

\[ s^2 = \frac{(37.4 - 46.42)^2 + (48.8 - 46.42)^2 + (46.9 - 46.42)^2 + (55.0 - 46.42)^2 + (44.0 - 46.42)^2}{(5-1)} \]

\[ = 41.682, \]

which gives \( s = 6.456 \).

(b) Since \( \sigma \) is unknown and \( n = 5 \) is small, we shall use \( t \) distribution to construct the desired confidence interval for \( \mu \). For \( \alpha = 0.10 \), we have \( t_{0.05}(4) = 2.132 \). Hence, a 90% confidence interval for \( \mu \) is given by

\[ \bar{x} \pm 2.132(s/\sqrt{5}) = 46.42 \pm 2.132(6.456/\sqrt{5}) = 46.42 \pm 6.16 \]

or \([40.26, 52.58]\).

4. A test was conducted to determine whether a wedge on the end of a plug fitting designed to hold a seal onto the plug was doing its job. The data taken were in the form of measurements of the force required to remove a seal from the plug with the
wedge in place (say, \(X\)) and the force required without the plug (say, \(Y\)). Assume that the distributions of \(X\) and \(Y\) are \(N(\mu_X, \sigma^2_X)\) and \(N(\mu_Y, \sigma^2_Y)\), respectively. Ten independent observations of \(X\) are

\[
3.26, 2.26, 2.62, 2.62, 2.36, 3.00, 2.62, 2.40, 2.30, 2.40.
\]

Ten independent observations of \(Y\) are

\[
1.80, 1.46, 1.54, 1.42, 1.32, 1.36, 1.64, 2.00, 1.54.
\]

(15) (a) Find \(\bar{x}, s^2_x, \bar{y}\) and \(s^2_y\).

(15) (b) Find a 95% confidence interval for \(\mu_X - \mu_Y\).

**Solution.** (a) Based on the provided sample data, we find

\[
\bar{x} = 2.584, s^2_x = 0.1042, \bar{y} = 1.564, s^2_y = 0.0428.
\]

(b) We first calculate \(s^2_p\):

\[
s^2_p = \frac{(n - 1)s^2_x + (m - 1)s^2_y}{n + m - 2} = \frac{(10 - 1)(0.1042) + (10 - 1)(0.0428)}{10 + 10 - 2} = 0.0735.
\]

It follows that \(s_p = \sqrt{0.0735} = 0.2711\). For \(1 - \alpha = 0.95\), i.e. \(\alpha = 0.05\), we have \(t_{\alpha/2}(n + m - 2) = t_{0.025}(18) = 2.101\). So, the 95% confidence interval for \(\mu_X - \mu_Y\) is

\[
[2.584 - 1.564 - (2.101)(0.2711)\sqrt{\frac{1}{10} + \frac{1}{10}}, 2.584 - 1.564 - (2.101)(0.2711)\sqrt{\frac{1}{10} + \frac{1}{10}}] = [0.7653, 1.2747].
\]