Homework #1
Due on Thursday, February 9, 2012

Instructor: Larry Ammann

CourseBook: John E. Freund’s Mathematical Statistics with Applications (7th Edition)
Contents

Exercise 8.2  
(a) .......................................................... 3  
(b) .......................................................... 3  
Exercise 8.3  .................................................. 4  
Exercise 8.21 ................................................ 4  
Exercise 8.22 ................................................ 5  
Exercise 8.23 ................................................ 6  
Exercise 8.24 ................................................ 6  
Exercise 8.31 ................................................ 7  
Exercise 8.37 ................................................ 7  
Exercise 8.64  
(a) .......................................................... 8  
(b) .......................................................... 8
Exercise 8.2

Use Theorem 4.14 and its corollary to show that if $X_{11}, X_{21}, \ldots, X_{1n_1}, X_{21}, X_{22}, \ldots, X_{2n_2}$ are independent random variables, with the first $n_1$ constituting a random sample from an infinite population with the mean $\mu_1$ and the variance $\sigma_1^2$ and the other $n_2$ constituting a random sample from an infinite population with the mean $\mu_2$ and the variance $\sigma_2^2$, then

(a) $E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2$;

Answer:

$$E(\bar{X}_1 - \bar{X}_2) = E\left( \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1i} - \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2i} \right)$$

$$= \frac{1}{n_1} \sum_{i=1}^{n_1} E(X_{1i}) - \frac{1}{n_2} \sum_{i=1}^{n_2} E(X_{2i})$$

$$= \frac{1}{n_1} \sum_{i=1}^{n_1} \mu_1 - \frac{1}{n_2} \sum_{i=1}^{n_2} \mu_2$$

$$= \frac{1}{n_1} n_1 \mu_1 - \frac{1}{n_2} n_2 \mu_2$$

$$= \mu_1 - \mu_2$$  \hspace{1cm} (a)

(b) $\text{var}(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

Answer: If $X_1, \ldots, X_n$ are independent random variables, then

$$\text{var}(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 \text{var}(X_i) \quad (1)$$

Since $\bar{X}_1$ and $\bar{X}_2$ are independent,

$$\text{var}(\bar{X}_1 - \bar{X}_2) = (1)^2 \text{var}(\bar{X}_1) + (-1)^2 \text{var}(\bar{X}_2)$$

$$= \text{var}(\bar{X}_1) + \text{var}(\bar{X}_2) \quad (2)$$

$$= \text{var}\left( \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1i} \right) + \text{var}\left( \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2i} \right) \quad (3)$$

$$= \frac{(1)^2}{n_1} \sum_{i=1}^{n_1} \text{var}(X_{1i}) + \frac{(-1)^2}{n_2} \sum_{i=1}^{n_2} \text{var}(X_{2i}) \quad (4)$$

$$= \frac{1}{n_1^2} \sum_{i=1}^{n_1} \text{var}(X_{1i}) + \frac{1}{n_2^2} \sum_{i=1}^{n_2} \text{var}(X_{2i}) \quad (5)$$

$$= \frac{1}{n_1^2} n_1 \sigma_1^2 + \frac{1}{n_2^2} n_2 \sigma_2^2 \quad (6)$$

$$= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \quad (7)$$

Note that (4) is obtained from (3) using the equality (1).
Exercise 8.3

With reference to Exercise 8.2, show that if the two samples come from normal populations, then $\overline{X}_1 - \overline{X}_2$ is a random variable having a normal distribution with the mean $\mu_1 - \mu_2$ and the variance $\frac{\sigma^2_1}{n_1} + \frac{\sigma^2_2}{n_2}$. (Hint: Proceed as in the proof of Theorem 8.4.)

**Answer:** Moment generating functions (MGF) of random variables $X_{1i}, X_{2i}, \ldots, X_{ni}$ are

$$M_{X_{1i}}(\theta) = e^{\mu_1 \theta + \frac{1}{2} \sigma^2_1 \theta^2}, i = \{1, 2, \ldots, n_1\}$$

Similarly, MGF of random variables $X_{11}, X_{12}, \ldots, X_{1n_1}$ are

$$M_{X_{1i}}(\theta) = e^{\mu_1 \theta + \frac{1}{2} \sigma^2_1 \theta^2}, i = \{1, 2, \ldots, n_2\}$$

Then, MGF of $\overline{X}_1$ and $\overline{X}_2$ are

$$M_{\overline{X}_1}(\theta) = e^{\mu_1 \theta + \frac{1}{2} \frac{\sigma^2_1}{n_1} \theta^2}, M_{\overline{X}_2}(\theta) = e^{\mu_2 \theta + \frac{1}{2} \frac{\sigma^2_2}{n_2} \theta^2}$$

**Theorem 1:** If $X_1, X_2, \ldots, X_n$ is a sequence of independent (and not necessarily identically distributed) random variables, and

$$S_n = \sum_{i=1}^n a_i X_i$$

where the $a_i$ are constants, then the probability density function for $S_n$ is the convolution of the probability density functions of each of the $X_i$, and the moment-generating function for $S_n$ is given by

$$M_{S_n}(t) = M_{X_1}(a_1 t) M_{X_2}(a_2 t) \cdots M_{X_n}(a_n t).$$

Using the above theorem:

$$M_{\overline{X}_1 - \overline{X}_2}(\theta) = M_{\overline{X}_1}(\theta) M_{\overline{X}_2}(-\theta)$$

$$= e^{\mu_1 \theta + \frac{1}{2} \frac{\sigma^2_1}{n_1} \theta^2} e^{-\mu_2 \theta + \frac{1}{2} \frac{\sigma^2_2}{n_2} \theta^2}$$

$$= e^{(\mu_1 - \mu_2) \theta + \frac{1}{2} \left(\frac{\sigma^2_1}{n_1} + \frac{\sigma^2_2}{n_2}\right) \theta^2}$$

Therefore, $\overline{X}_1 - \overline{X}_2$ has a normal distribution with mean $\mu_1 - \mu_2$ and variance $\frac{\sigma^2_1}{n_1} + \frac{\sigma^2_2}{n_2}$.

Exercise 8.21

Use Theorem 8.11 to show that, for random samples of size $n$ from a population with the variance $\sigma^2$, the sampling distribution of $S^2$ has mean $\sigma^2$ and the variance $\frac{2\sigma^4}{n-1}$.
Answer: Let $Y$ be $\frac{(n-1)S^2}{\sigma^2}$. Then by theorem 8.11, $Y$ has a chi-square distribution with $n-1$ degrees of freedom. We know that chi-square distribution with $n-1$ degrees of freedom has the mean $n-1$ and variance $2(n-1)$. Let’s write the sampling distribution in terms of $Y$ which has a known mean and variance.

$$S^2 = \frac{\sigma^2}{n-1} Y$$

(8)

By using the linearity of expectation,

$$E[S^2] = \frac{\sigma^2}{n-1} E[Y]$$

$$= \frac{\sigma^2}{n-1} (n-1)$$

$$= \sigma^2$$

Similarly, for variance,

$$\text{var}(S^2) = \left( \frac{\sigma^2}{n-1} \right)^2 \text{var}(Y)$$

$$= \frac{\sigma^4}{(n-1)^2} 2(n-1)$$

$$= \frac{2\sigma^4}{n-1}$$

Exercise 8.22

Show that if $X_1, X_2, \ldots, X_n$ are independent random variables having a chi-square distribution with $\nu = 1$ and $Y_n = X_1 + X_2 + \ldots + X_n$, then the distribution of

$$Z = \frac{Y_n n - 1}{\sqrt{2/n}}$$

(9)

as $n \to \infty$ is the standard normal distribution.
Answer:

\[
Z = \frac{Y_n - 1}{\sqrt{2/n}} \quad (10)
\]
\[
= \frac{Y_n - n}{\sqrt{2\sqrt{n}}} \quad (11)
\]

Chi-square distribution with 1 degree of freedom has \(\mu = 1\) and \(\sigma^2 = 2\) (hence \(\sigma = \sqrt{2}\)). \(Y_n\) is the sum of i.i.d random variables with \(\mu = 1\) and \(\sigma = \sqrt{2}\). Hence by central limit theorem (as shown below)

\[
\lim_{n \to \infty} P\left(\frac{Y_n - n}{\sqrt{2\sqrt{n}}} \leq x\right) = \lim_{n \to \infty} P(Z \leq x) = \Phi(x) \quad (13)
\]

Central Limit Theorem: Let \(X_1, X_2, \ldots, X_n\) be i.i.d random variables having mean \(\mu\) and finite nonzero variance \(\sigma^2\). Let \(Y_n = X_1 + X_2 + \ldots + X_n\). Then

\[
\lim_{n \to \infty} P\left(\frac{Y_n - n}{\sigma\sqrt{n}} \leq x\right) = \Phi(x) \quad (14)
\]

where \(\Phi(x)\) is the probability that a standard normal random variable is less than \(x\). Note that (11) Therefore, by CTL, \(Z\) is the standard normal distribution as \(n \to \infty\).

Exercise 8.23

Based on the result of Exercise 8.23, show that if \(X\) is a random variable having chi-square distribution with \(v\) degrees of freedom and \(v\) is large, then distribution of \(\frac{X - v}{\sqrt{2v}}\) can be approximated with the standard normal distribution.

Answer: In the previous question \(Y_n\) is the sum of i.i.d random variables with chi-square distribution with \(\mu = 1\) and \(\sigma^2 = 2\). The sum of chi-square distributions has also a chi-square distribution with degree as the sum of the degrees of each random variable. Therefore, if \(X_1, X_2, \ldots, X_v\) are i.i.d random variables having chi-square distribution with 1 degrees of freedom and \(X = X_1 + X_2 + \ldots + X_v\), then \(X\) has a chi-square distribution with \(v\) degrees of freedom. We also know that \(X\) has the mean \(v\) and variance \(2v\). By central limit theorem, \(\frac{X - v}{\sqrt{2v}}\) is approximately \(N(0,1)\).

Exercise 8.24

Use the method of Exercise 8.23 to find the approximate value of that random variable having a chi-square distribution with \(v = 50\) having a value greater than 68.

Answer: Let \(Y\) be the variable having a chi-square distribution with 50 degrees of freedom. Then,

\[
P(Y > 68) = P\left(\frac{Y - 50}{\sqrt{100}} \geq \frac{68 - 50}{1\sqrt{100}}\right) \quad (15)
\]
\[
\approx P(Z > 1.8) \quad (16)
\]
\[
= 1 - \Phi(1.8) \quad (17)
\]
\[
= 1 - 0.9641 \quad (18)
\]
\[
= 0.0359 \quad (19)
\]
Exercise 8.31

Show that for $V > 2$ the variance of the $t$ distribution with $v$ degrees of freedom is $\frac{v}{v-2}$.

**Answer:** We can use the following variance formula: $\text{var}(X) = E[X^2] - (E[X])^2$. Let’s first compute $E[X^2]$.

\[
E[X^2] = \int_{-\infty}^{\infty} t^2 f_X(t) \, dt = \int_{-\infty}^{0} t^2 f_X(t) \, dt + \int_{0}^{\infty} t^2 f_X(t) \, dt \tag{20}
\]

\[
= -\int_{\infty}^{0} t^2 f_X(-t) \, dt + \int_{0}^{\infty} t^2 f_X(t) \, dt \tag{21}
\]

\[
= 2 \int_{0}^{\infty} t^2 f_X(t) \, dt \tag{22}
\]

\[
= 2c \int_{0}^{\infty} t^2 (1 + \frac{t^2}{v})^{-\frac{v+1}{2}} \, dt \tag{23}
\]

\[
= 2c \int_{0}^{\infty} vu(1 + u)^{-v/2 - 1/2} \frac{\sqrt{v}}{2\sqrt{u}} \, du \tag{24}
\]

\[
= cn^{3/2} \int_{0}^{\infty} u^{3/2 - 1}(1 + u)^{-3/2 - (v/2 - 1)} \, du \tag{25}
\]

\[
= cn^{3/2}B(3/2, v/2 - 1) \tag{26}
\]

\[
= \frac{1}{\sqrt{v} B(\frac{3}{2}, \frac{v}{2} - 1)} \tag{27}
\]

\[
= \frac{\Gamma(v/2 + 1/2) \Gamma(1/2 + 1)\Gamma(v/2 - 1)}{\Gamma(v/2)\Gamma(1/2)} \tag{28}
\]

\[
= n^n(n/2)\Gamma(1/2) \tag{29}
\]

\[
= \frac{\Gamma(1/2 + 1)\Gamma(v/2 - 1)}{\Gamma(v/2)\Gamma(1/2)} \tag{30}
\]

\[
= \frac{\Gamma(1/2)\Gamma(v/2)\Gamma(v/2)\Gamma(1/2)}{\Gamma(v/2)\Gamma(1/2)} \tag{31}
\]

\[
= \frac{v}{v - 2} \tag{32}
\]

Exercise 8.37

Verify that if $X$ has an $F$ distribution with $\nu_1$ and $\nu_2$ degree of freedom and $\nu_2 \to \infty$, the distribution of $Y = \nu_1 X$ approaches the chi-square distribution with $\nu_1$ degrees of freedom.
**Answer:** Note that for large enough degrees of freedom \( v \),

\[
\frac{v s^2}{\sigma^2} \approx \chi_v^2
\]  
(33)

By Weak Law of Large Numbers

\[
s^2 = \frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x})^2 \overset{p}{\to} \sigma^2
\]  
(34)

combining this with Slutsky’s theorem, we get

\[
\frac{s^2}{\sigma^2} \overset{p}{\to} 1
\]

Limiting distribution of chi-square divided it by its degree of freedom converges to 1 in probability by Slutsky’s theorem (by (33) and (34)). Therefore, \( X \) will have approximately a chi-square distribution divided by its degree of freedom \( v_1 \). Then, \( Y = v_1 X \) will have approximately a chi-square distribution, by Slutsky’s theorem as shown below.

Slutsky’s theorem:

\[
X_n \overset{d}{\to} X, Y_n \overset{d}{\to} c
\]

Then,

(1) \( X_n Y_n \overset{d}{\to} cX \)

---

**Exercise 8.64**

A random sample of size \( n = 81 \) is taken from an infinite population with the mean \( \mu = 128 \) and the standard deviation \( \sigma = 6.3 \). With what probability can we assert that the value we obtain for \( \bar{X} \) will not fall between 126.6 and 129.4?

(a)

Chebyshev’s theorem:

**Answer:** Chebyshev’s theorem states that

\[
\Pr(|\bar{X} - \mu| \geq k\sigma) \leq \frac{1}{k^2}.
\]

(\text{a})

However, note that Chebyshev’s inequality does not give useful information when \( k < 1 \). In our case, \( k = \frac{1.6}{6.3} < 1 \), so we can not use Chebyshev’s theorem.

(b)

the central limit theorem?
**Answer:** Central limit theorem states that if a random sample of $n$ observations is selected from a population (any population), then when $n$ is sufficiently large, the sampling distribution of $X$ will be approximately normal. (The larger the sample size, the better will be the normal approximation to the sampling distribution of $X$.) Let $Z = \frac{X - 128}{6.3/\sqrt{81}}$. By CLT, $Z$ has approximately standard normal distribution.

\[
P(126.6 < X < 129.4) = P(126.6 - 128 < X - 128 < 129.4 - 128) = P(-1.4 < X - 128 < 1.4) = P\left(\frac{-1.4}{6.3/\sqrt{81}} < \frac{X - 128}{6.3/\sqrt{81}} < \frac{1.4}{6.3/\sqrt{81}}\right) = P(-2 < Z < 2) = \Phi(2) - \Phi(-2) = 0.9772 - 0.0228 = 0.9544
\]

Question asks the probability that value will not fall between 126.6 and 129.4. So the answer is $1 - 0.9544 = 0.0456$. 

(b)