Tutorial Letter 203/2/2013

Distribution theory II

STA2603

Semester 2

Department of Statistics

Solutions to Assignment 03

BAR CODE

QUESTION 1

We assume that the cumulative distribution function of X is as follows:

$$F_X(k) = \begin{cases} 0 & \text{for } k < 1 \\ 0.1 & \text{for } 1 \leq k < 2 \\ 0.3 & \text{for } 2 \leq k < 3 \\ 0.7 & \text{for } 3 \leq k < 4 \\ 0.8 & 4 \leq k < 5 \\ 1 & k \geq 5 \end{cases}$$

(a) Investigating the jumps of the function F_X, we find that the frequency function (probability mass function) is

$$p_X(k) = \begin{cases} 0.1 & \text{for } k = 1 \\ 0.2 & \text{for } k = 2 \\ 0.4 & \text{for } k = 3 \\ 0.1 & \text{for } k = 4 \\ 0.2 & \text{for } k = 5 \end{cases}$$

(i)

$$\begin{aligned} E(X) &= \sum_k k \cdot p_X(k) \\ &= 1 \cdot 0.1 + 2 \cdot 0.2 + 3 \cdot 0.4 + 4 \cdot 0.1 + 5 \cdot 0.2 \\ &= 3.1 \end{aligned}$$

(ii)

$$\begin{aligned} E(\frac{1}{X}) &= \sum_k \frac{1}{k} \cdot p_X(k) \\ &= \frac{1}{1} \cdot 0.1 + \frac{1}{2} \cdot 0.2 + \frac{1}{3} \cdot 0.4 + \frac{1}{4} \cdot 0.1 + \frac{1}{5} \cdot 0.2 \\ &= 0.39833 \end{aligned}$$

(iii)

$$\begin{aligned} E(X^2 - 1) &= \sum_k \left(k^2 - 1\right) \cdot p_X(k) \\ &= \left(1^2 - 1\right) \cdot 0.1 + \left(2^2 - 1\right) \cdot 0.2 + \left(3^2 - 1\right) \cdot 0.4 + \left(4^2 - 1\right) \cdot 0.1 + \left(5^2 - 1\right) \cdot 0.2 \\ &= 10.1. \end{aligned}$$

Alternatively, we can calculate this as $E(X^2-1) = E\left(X^2\right) - 1$ where

$$
\begin{aligned}
E(X^2) &= \sum_k \left(k^2\right) \cdot p_X(k) \\
&= \left(1^2\right) \cdot 0.1 + \left(2^2\right) \cdot 0.2 + \left(3^2\right) \cdot 0.4 + \left(4^2\right) \cdot 0.1 + \left(5^2\right) \cdot 0.2 \\
&= 11.1
\end{aligned}
$$

and therefore $E(X^2-1) = E\left(X^2\right) - 1 = 11.1 - 1 = 10.1.$

(b)

$$Var(X) = E\left(X^2\right) - E\left(X\right)^2$$

where $E\left(X^2\right) = 11.1$ can be found from (a) (iii) (using $E\left(X^2\right) = E\left(X^2-1\right) + 1$ if you found $E\left(X^2-1\right)$ directly) and $E\left(X\right) = 3.1$was found in (a) (i). Therefore,

$$Var(X) = 11.1 - (3.1)^2 = 1.49.$$

QUESTION 2

If the joint density of U and V is

$$f_{U,V}(u,v) = \begin{cases} 6(1-u-v) & \text{for } 0 < u < 1, 0 < v < 1-u \\ 0 & \text{elsewhere} \end{cases}$$

then to find $E\left(U\right)$ we can either integrate directly or we can find the marginal density of U first. The first method gives

$$
\begin{aligned}
E\left(U\right) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u f_{U,V}(u,v)dudv \\
&= \int_0^1 \left(\int_0^{1-u} u \cdot 6(1-u-v)dv\right) du \\
&= \int_0^1 6u \left(\int_0^{1-u} (1-u-v)dv\right) du \\
&= \int_0^1 6u \left(\left(v - uv - \frac{1}{2}v^2\right)\Big|_0^{1-u}\right) du \\
&= \int_0^1 6u \left(\left(1-u-u\left(1-u\right) - \frac{1}{2}\left(1-u\right)^2\right)\right) du \\
&= 3\int_0^1 \left(u - 2u^2 + u^3\right) du = 3\left(\frac{1}{2}u^2 - \frac{2}{3}u^3 + \frac{1}{4}u^4\right)\Big|_0^1 \\
&= \frac{1}{4}.
\end{aligned}
$$

To apply the second method, we first find

$$f_U(u) = \int_{-\infty}^{\infty} f_{U,V}(u,v)dv$$

which for $0 < u < 1$ gives

$$f_U(u) = \int_0^{1-u} 6(1-u-v)dv = \ldots = 3(u-1)^2.$$

Therefore,

$$E(U) = \int_{-\infty}^{\infty} u f_U(u)du = \int_0^1 u3(u-1)^2 du = \frac{1}{4}.$$

Similarly, we get

$$\begin{aligned} E(V) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} v f_{U,V}(u,v)dudv = \int_0^1 \left(\int_0^{1-u} v \cdot 6(1-u-v)dv\right) du \\ &= \ldots = \frac{1}{4}. \end{aligned}$$

[Note that the joint density of U and V and identically be written as

$$f_{U,V}(u,v) = \begin{cases} 6(1-u-v) & \text{for } 0 < v < 1, 0 < u < 1-v \\ 0 & \text{elsewhere} \end{cases}$$

Therefore, since the density function stays the same with u and v swapped, the marginal densities of U and V are the same, and in particular $E(U) = E(V)$.]

Finally,

$$Cov(U,V) = E(UV) - E(U)E(V)$$

where

$$\begin{aligned} E(V) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} uv f_{U,V}(u,v)dudv = \int_0^1 \left(\int_0^{1-u} uv \cdot 6(1-u-v)dv\right) du \\ &= \ldots = \frac{1}{20} \end{aligned}$$

and therefore

$$Cov(U,V) = \frac{1}{20} - \frac{1}{4}\cdot\frac{1}{4} = -\frac{1}{80}.$$

QUESTION 3

(a) The frequency function of the binomial distribution with parameters n and p is

$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 2, \ldots, n.$$

Therefore, the moment generating function is

$$\begin{aligned} M_X(t) &= E\left(e^{Xt}\right) = \sum_{x=0}^{n} e^{xt} p_X(x) = \sum_{x=0}^{n} e^{xt} \binom{n}{x} p^x (1-p)^{n-x} \\ &= \sum_{x=0}^{n} \binom{n}{x} \left(pe^t\right)^x (1-p)^{n-x}. \end{aligned}$$

Now, applying the formula for the finite binomial series, see page 17 of the study guide, we identify the sum on the right as being of the type

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$$

with $a = pe^t$, $b = 1-p$. Therefore we conclude that

$$\begin{aligned} M_X(t) &= \left(pe^t + (1-p)\right)^n \\ &= \left(0.2e^t + 0.8\right)^{50}. \end{aligned}$$

(b) The first two moments of X can be found by differentiating the moment generating function with respect to t:

$$\begin{aligned} E(X) &= M_X'(0), \\ E\left(X^2\right) &= M_X''(0). \end{aligned}$$

With

$$M_X(t) = \left(0.2e^t + 0.8\right)^{50},$$

we get (using the product and chain rules of differentiation)

$$\begin{aligned} M_X'(t) &= 50\left(0.2e^t + 0.8\right)^{49} \cdot 0.2e^t = 10e^t\left(0.2e^t + 0.8\right)^{49}, \\ M_X''(t) &= 49 \cdot 10e^t\left(0.2e^t + 0.8\right)^{48} \cdot 0.2e^t + 10e^t\left(0.2e^t + 0.8\right)^{49} \\ &= 98e^{2t}\left(0.2e^t + 0.8\right)^{48} + 10e^t\left(0.2e^t + 0.8\right)^{49} \end{aligned}$$

and therefore,

$$\begin{aligned} E(X) &= M_X'(0) = 10e^0\left(0.2e^0 + 0.8\right)^{49} = 10, \\ E\left(X^2\right) &= M_X''(0) = 98e^{2\cdot 0}\left(0.2e^0 + 0.8\right)^{48} + 10e^0\left(0.2e^0 + 0.8\right)^{49} = 108, \\ Var(X) &= E\left(X^2\right) - (E(X))^2 = 108 - 100 = 8. \end{aligned}$$

QUESTION 4

(a) If the density function of X is

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x > 0 \\ 0 & \text{for } x < 0 \end{cases}$$

then the moment generating function of X is given by

$$\begin{aligned} M_X(t) &= E\left(e^{tX}\right) = \int_{-\infty}^{\infty} e^{tx} f_x(x) dx = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx \\ &= \lambda \int_0^{\infty} e^{(t-\lambda)x} dx = \frac{\lambda}{t-\lambda} \left(e^{(t-\lambda)x}\right)\Big|_0^{\infty} = \frac{\lambda}{t-\lambda}(0-1) \\ &= \frac{\lambda}{\lambda - t} = \left(1 - \frac{t}{\lambda}\right)^{-t}. \end{aligned}$$

The integral only has a finite value if $t < \lambda$. R

(b) Differentiating the moment generating function of X we get

$$\begin{aligned} M_X(t) &= \lambda(\lambda - t)^{-1} \\ M_X'(t) &= \lambda(\lambda - t)^{-2} \\ M_X''(t) &= 2\lambda(\lambda - t)^{-3} \\ M_X'''(t) &= 6\lambda(\lambda - t)^{-4} \end{aligned}$$

and therefore

(i)

$$E(X) = M_X'(0) = \lambda(\lambda - 0)^{-2} = \frac{1}{\lambda}$$

(ii)

$$\begin{aligned} E(X^2) &= M_X''(0) = 2\lambda(\lambda - 0)^{-3} = \frac{2}{\lambda^2}, \\ Var(X) &= E(X^2) - (E(X))^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \end{aligned}$$

(iii)

$$E(X^3) = M_X'''(0) = 6\lambda(\lambda - 0)^{-4} = \frac{6}{\lambda^3}.$$

QUESTION 5

Let X be the first time that a defective drill is found. Then X has the geometric distribution with $p = 0.05$. The probability distribution function is

$$p_X(x) = \begin{cases} (1-p)^{x-1}\,p & \text{for } x = 1, 2, \ldots \\ 0 & \text{otherwise.} \end{cases}$$

The expected value is therefore

$$E(X) = \sum_{x=1}^{\infty} x\,(1-p)^{x-1}\,p = \frac{1}{p}.$$

(See Example B on page 117 of the textbook so see how this sum can be evaluated!)
In our case $p = 0.05$,so we see that on average $1/0.05 = 20$ drills need to be inspected before finding the first defective one.

QUESTION 6

(a) If X has the density function

$$f(x) = \begin{cases} e^x & \text{for } x \leq 0 \\ 0 & \text{elsewhere} \end{cases}$$

then

$$E\left(e^{\frac{3X}{2}}\right) = \int_{-\infty}^{0} e^{\frac{3}{2}x} e^x\,dx = \frac{2}{5}.$$

(b) The moment generating function of X is

$$\begin{aligned} M(t) &= E\left(e^{tX}\right) = \int_{-\infty}^{0} e^{tx} e^x\,dx = \int_{-\infty}^{0} e^{(t+1)x}\,dx \\ &= \frac{1}{t+1}\int_{-\infty}^{0} (t+1)\,e^{(t+1)x}\,dx = \frac{1}{t+1}\left(e^{(t+1)x}\right)\Big|_{-\infty}^{0} \\ &= \frac{1}{t+1}, \quad \text{as long as } t > -1. \end{aligned}$$

(c) Since we have the moment generating function, the easiest way to find $Var(X)$ is by using the moment generating function to first find $E(X)$ and $E\left(X^2\right)$ .Since we can write $M(t) = (t+1)^{-1}$, the first two derivatives of M are

$$\begin{aligned} M'(t) &= -(t+1)^{-2}, \\ M''(t) &= 2(t+1)^{-3} \end{aligned}$$

and therefore

$$\begin{aligned} E(X) &= M'(0) = -(0+1)^{-2} = -1, \\ E(X^2) &= M''_X(0) = 2\,(0+1)^{-3} = 2, \\ Var\,(X) &= E(X^2) - (E\,(X))^2 = 2 - (-1)^2 = 1. \end{aligned}$$

QUESTION 7

According to the central limit theorem, $\bar{X}$ has approximately the normal distribution with expected value μ and variance σ^2/n, meaning $\left(\bar{X}-\mu\right)/\left(\sigma/\sqrt{n}\right)$ has approximately the standard normal distribution. Here, μ and σ^2 are the mean and variance of the gamma distribution with $\alpha = 2$ and $\lambda = 4$ and $n = 128$. So, here we have

$$\mu = \frac{\alpha}{\lambda} = \frac{1}{2}, \quad \sigma^2 = \frac{\alpha}{\lambda^2} = \frac{1}{8}, \quad \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{1/8}}{\sqrt{128}} = \frac{1}{32}.$$

Therefore the requested probability is found as

$$\begin{aligned} P(7 < \overline{X} < 9) = P\left(\frac{7-\frac{1}{2}}{\frac{1}{32}} < \frac{\bar{X}-\frac{1}{2}}{\frac{1}{32}} < \frac{9-\frac{1}{2}}{\frac{1}{32}}\right) \\ \approx P\left(\frac{7-\frac{1}{2}}{\frac{1}{32}} < Z < \frac{9-\frac{1}{2}}{\frac{1}{32}}\right) \approx P\,(208.0 < Z < 272.0) \approx 0. \end{aligned}$$

QUESTION 8

Here $X_1, X_2, \cdots, X_{12}$ denote a random sample from a $N(5,9)$ distribution, $Y_1, Y_2, \cdots, Y_{10}$ an independent random sample from a $N(5,25)$ distribution, and

$$\overline{X} = \frac{1}{12}\sum_{i=1}^{12} X_i \qquad \overline{Y} = \frac{1}{10}\sum_{j=1}^{10} Y_j$$

$$S_X^2 = \frac{1}{11}\sum_{i=1}^{12}(X_i - \overline{X})^2 \qquad S_Y^2 = \frac{1}{9}\sum_{j=1}^{10}(Y_j - \overline{Y})^2$$

$$\hat{\sigma}_X^2 = \frac{1}{12}\sum_{i=1}^{12}(X_i - 5)^2 \qquad \hat{\sigma}_Y^2 = \frac{1}{10}\sum_{j=1}^{10}(Y_j - 5)^2$$

(a) Then

(i) $T = \dfrac{X_i - 5}{3}$ follows a standard normal distribution (mean 0, variance 1).

(ii) $Q = \sum_{i=2}^{12}\left(\dfrac{X_i - 5}{3}\right)^2$ follows a chi-square distribution with parameter (degree of freedom) 11. This is because each $\dfrac{X_i - 5}{3}$ is a standard normal variable, $\left(\dfrac{X_i - 5}{3}\right)^2$ has the $\chi^2(1)$ distribution, and as the sum of 11 independent chi-square variables with degree 1, Q is chi-square variable with 11 degrees of freedom.

(iii) $R = 12\left(\dfrac{\overline{X} - 5}{3}\right)^2$ follows a chi-square distribution with parameter (degree of freedom) one. This is because

$$R = 12\left(\frac{\overline{X} - 5}{3}\right)^2 = \left(\frac{\frac{1}{12}\sum_{i=1}^{12} X_i - 5}{3/\sqrt{12}}\right)^2$$

is the square of a standard normal random variable.

(iv) $U = \sum_{i=1}^{7}\left(\dfrac{X_i - 5}{3}\right)^2$ follows a chi-square distribution with parameter (degree of freedom) seven, since it is the sum of seven squares of independent standard normal random variables.

(v) $V = \dfrac{U/7}{Q/11}$ does not follows an $F(7, 11)$ distribution since U and Q are not independent of each other.

(iv) $W = \dfrac{25\hat{\sigma}^2_X}{9\hat{\sigma}^2_Y}$ follows an $F(12, 10)$ distribution, since

$$W = \frac{25\hat{\sigma}^2_X}{9\hat{\sigma}^2_Y} = \frac{25 \cdot \frac{1}{12}\sum_{i=1}^{12}(X_i - 5)^2}{9 \cdot \frac{1}{10}\sum_{j=1}^{10}(Y_j - 5)^2} = \frac{\sum_{i=1}^{12}(\frac{X_i - 5}{3})^2/12}{\sum_{j=1}^{10}(\frac{Y_j - 5}{5})^2/10}$$

where the sum in the sum $\sum_{i=1}^{12}(\frac{X_i-5}{3})^2$ is a $\chi^2\ (12)$ random variable and the sum $\sum_{j=1}^{10}(\frac{Y_j-5}{5})^2$ is a $\chi^2\ (10)$ random variable, and the two are independent of each other.

(b) As found in (a) (iv), $W = \dfrac{25\hat{\sigma}_X^2}{9\hat{\sigma}_Y^2}$ has the $F\,(12, 10)$ distribution, and therefore $E\,(W) = \frac{10}{10-8} = 5$, and therefore

$$E\left(\frac{\hat{\sigma}_X^2}{\hat{\sigma}_Y^2}\right) = \frac{9}{25}E\,(W) = \frac{9}{5} = 1.8.$$

(c) Since $\overline{X}$ has the normal distribution with mean 5 and variance $9/12$, and $\overline{Y}$ has the normal distribution with mean 5 and variance 25/10, and the two random variables are independent of each other, it follows that $\overline{X} - \overline{Y}$ has the normal distribution with mean 0 and variance $\frac{9}{12} + \frac{25}{10} = \frac{13}{4}$. Therefore, since $\left(\overline{X} - \overline{Y}\right)/\sqrt{13/4}$ is a standard normal random variable,

$$P(\overline{X} - \overline{Y} \ > \ 2) = P(\frac{\overline{X} - \overline{Y}}{\sqrt{13/4}} > \frac{2}{\sqrt{13/4}}) = P\left(Z > \frac{2}{\sqrt{13/4}}\right)$$

$$\approx \ P\,(Z > 1.1094) \approx 1 - 0.8643 = 0.1357.$$