

STAT7203: Applied Probability and Statistics  
Assignment 2

Due by 11:00 am on Tuesday the 24th of September, 2019  
via the Electronic Assignment Submission System (62-225)

The marks for each question is indicate by the number in square brackets. There are a total of 10 marks for this assignment.

1. Suppose that Let  $Z_1$  and  $Z_2$  be two independent standard normal random variables. For constants  $a_1, a_2, b_{11}, b_{12}, b_{21}, b_{22}$ , define the random variables  $X$  and  $Y$  such that

$$\begin{aligned}X &= a_1 + b_{11}Z_1 + b_{12}Z_2 \\Y &= a_2 + b_{21}Z_1 + b_{22}Z_2.\end{aligned}$$

- (a) Determine those constants which result in

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathbf{N} \left( \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix} \right).$$

[2]

*Solution:*

$$\begin{aligned}\mathbb{E}[X] &= a_1 = 2 \\ \mathbb{E}[Y] &= a_2 = -1 \\ \text{Var}(X) &= b_{11}^2 + b_{12}^2 = 3 \\ \text{Var}(Y) &= b_{21}^2 + b_{22}^2 = 2 \\ \text{Cov}(X, Y) &= b_{11}b_{21} + b_{12}b_{22} = -1\end{aligned}$$

Various solutions for the  $b$ 's are possible. One possible solution is to set  $b_{12} = 0$  from which we get  $b_{11} = \sqrt{3}$ ,  $b_{21} = -1/\sqrt{3}$  and  $b_{22} = \sqrt{5/3}$ .

- (b) Determine the probability that  $Y \geq 0.5$ .

[1]

*Solution:* The marginal distribution of  $Y$  is  $\mathbf{N}(-1, 2)$ . So

$$\mathbb{P}(Y \geq 0.5) = \mathbb{P}\left(\frac{Y - -1}{\sqrt{2}} \geq \frac{0.5 - -1}{\sqrt{2}}\right) = \mathbb{P}(Z \geq 1.5/\sqrt{2}) = 0.1444222$$

2. A continuous random variable  $X$  has probability density function

$$f_X(x) = \frac{1}{2}e^{-|x|}, \quad x \in \mathbb{R}.$$

- (a) Find the moment generating function  $M_X(t)$  of  $X$ , remembering to state the valid range for  $t$ .

[2]

*Solution:*

$$\begin{aligned}\mathbb{E}[e^{tX}] &= \int_{-\infty}^{\infty} e^{tx} \frac{1}{2}e^{-|x|} dx \\ &= \int_{-\infty}^0 \frac{1}{2}e^{(t+1)x} dx + \int_0^{\infty} \frac{1}{2}e^{(t-1)x} dx \\ &= \frac{1}{2}(t+1)^{-1} + \frac{1}{2}(t-1)^{-1} = (1-t^2)^{-1}, \quad |t| < 1.\end{aligned}$$

- (b) Let  $X_1, \dots, X_n$  be a simple random sample where  $X_i$  has the probability density function  $f_X$  given above and define  $\bar{X} = n^{-1}(X_1 + X_2 + \dots + X_n)$ . Using Markov's inequality it can be shown that for any  $a \in \mathbb{R}$  and any  $t \in \mathbb{R}$  for which the expectations below exist

$$\mathbb{P}(\bar{X} \geq a) = \mathbb{P}(e^{\bar{X}t} \geq e^{at}) \leq \frac{\mathbb{E}(e^{t\bar{X}})}{e^{at}}.$$

Define  $\mathcal{H}(t; a) := e^{-at} \mathbb{E}(e^{t\bar{X}})$ . For a fixed value of  $a$ , find the value  $t_a$  which minimises  $\mathcal{H}(t; a)$ . [3]

*Solution:* We first need the moment generating function of  $\bar{X}$ .

$$\begin{aligned} \mathbb{E}[e^{t\bar{X}}] &= \mathbb{E}[e^{t(X_1 + \dots + X_n)/n}] = \mathbb{E}[e^{tX_1/n}] \dots \mathbb{E}[e^{tX_n/n}] \\ &= (1 - (t/n)^2)^{-n}, \quad |t| < n. \end{aligned}$$

So we want to minimise  $\mathcal{H}(t; a) = e^{-at}(1 - (t/n)^2)^{-n}$ . Noting that this is equivalent to minimising  $\log \mathcal{H}(t; a)$ ,

$$\begin{aligned} \log \mathcal{H}(t; a) &= -at - n \log(1 - (t/n)^2) \\ \frac{d}{dt} \log \mathcal{H}(t; a) &= -a + \frac{2(t/n)}{1 - (t/n)^2} = -a + \frac{2tn}{n^2 - t^2} \end{aligned}$$

Setting the derivative equal to zero, we see the stationary points are given by solutions of

$$at^2 + 2nt - an^2 = 0$$

Solving the quadratic equation gives,  $t = \frac{n}{a}(-1 \pm \sqrt{1 + a^2})$ . The second derivative of  $\log \mathcal{H}(t; a)$  is

$$\frac{d^2}{dt^2} \log \mathcal{H}(t; a) = \frac{2n(n^2 + t^2)}{(n^2 - t^2)^2} > 0$$

for all  $t$  in the valid range. For  $a < 1$ , the negative solution will be outside the valid range for  $t$  so the positive solution  $t_a = \frac{n}{a}(\sqrt{1 + a^2} - 1)$  minimises  $\mathcal{H}(t; a)$ .

- (c) Let  $n = 100$  and  $a = 0.1$ . Compute the bound on  $\mathbb{P}(\bar{X} \geq 0.1)$  and compare this with the approximation from the central limit theorem. [2]

*Solution:* The bound from part (b) gives  $\mathbb{P}(\bar{X} \geq 0.1) \leq 0.7790434$ . To apply the CLT we need the mean and variance of  $\bar{X}$ . We can use the moment generating function

$$\begin{aligned} M_X(t) &= (1 - t^2)^{-1} \\ M'_X(t) &= 2t(1 - t^2)^{-2} & M'_X(0) &= 0 \\ M''_X(t) &= 2(1 - t^2)^{-2} + 8t^2(1 - t^2)^{-3} & M''_X(0) &= 2 \end{aligned}$$

So  $\mathbb{E}[\bar{X}] = 0$  and  $\text{Var}(\bar{X}) = 2/100$ . Therefore  $\mathbb{P}(\bar{X} \geq 0.1) = \mathbb{P}(Z \geq 0.1/\sqrt{0.02}) = \mathbb{P}(Z \geq 0.707) = 0.239$ . The bound constructed in part (b) is much larger than the approximation from the CLT for this value of  $a$ .