MAST90105: Lab and Workshop Problems for Week 6

The Lab and Workshop this week cover problems arising from Module 3, Section 4 (Normal Distribution and Central Limit Theorem), Module 4 and Module 5, Section 1.

1 Lab

1. Suppose $X_1 \stackrel{d}{=} N(\mu = 3, \sigma^2 = 4)$, $X_2 \stackrel{d}{=} N(3, 4)$, and X_1 and X_2 are independent.

   a. Let $Y = 5X_1 - 2X_2 + 6$. Name the distribution of Y and give the values of the associated parameters.

   b. Use the command `rnorm(n,mean,sd)` in R to generate a sample of size $n = 100$, $Y_1, \ldots, Y_{100}$, from this distribution. Use the command `mean()` to compute the sample mean and standardized second moment:

   $$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i, \quad \bar{Y}_S^2 = \frac{1}{n}\sum_{i=1}^{n} \frac{(Y_i - 15)^2}{116}.$$

   Repeat this procedure 1000 times to get 1000 sample means and sample second moments, $\bar{Y}_1, \ldots, \bar{Y}_{1000}$, and $\bar{Y}_{S,1}^2, \ldots, \bar{Y}_{S,1000}^2$.

   c. Estimate the mean and variance of $\bar{Y}$ and $\bar{Y}_S^2$ from the data $\bar{Y}_1, \ldots, \bar{Y}_{1000}$, and $\bar{Y}_{S,1}^2, \ldots, \bar{Y}_{S,1000}^2$. How close are these estimates to theoretical values?

   d. Standardize the sample means and sample second moments, $\bar{Y}_1, \ldots, \bar{Y}_{1000}$, and $\bar{Y}_{S,1}^2, \ldots, \bar{Y}_{S,1000}^2$, so that they have zero mean and unit variance. Use the command `hist()` to construct the histograms of the standardized data. What distribution do the histograms resemble? What are the exact distributions?

2. We would like to generate a random variable X which has the following cdf:

   $$F(x) = \begin{cases} 0, & x < 0, \\ x^3, & 0 \le x < 1, \\ 1, & x \ge 1. \end{cases}$$

   a. Use `runif(x)` function in R to generate random numbers from the uniform $U(0, 1)$ distribution and write a function that generates random numbers from the cdf $F(x)$.

   b. Use this function to generate a sample of size 1000 from this distribution and use this sample to estimate $\Pr(0.6 < X < 0.8)$. Compute the exact probability. How close is the estimated probability to the exact one?

   c. Use the generated sample to construct the histogram of this distribution. Comment on the shape of the histogram.

2 Workshop

You can use R for numerical calculations as needed.

3. The serum zinc level X in micrograms per deciliter for males between ages 15 and 17 has a distribution which is approximately normal with $\mu = 90$ and $\sigma = 15$. Compute the conditional probability $P(X > 120|X > 105)$.

4. Let $Z_1, Z_2, \cdots, Z_7$ be a random sample from the standard normal distribution $N(0, 1)$. Let $W = Z_1^2 + Z_2^2 + \cdots + Z_7^2$. Name the distribution of W with associated parameter value. Justify your answer.

5. If $X_1, X_2, \cdots, X_{16}$ is a random sample of size $n = 16$ from the normal distribution $N(50, 100)$.

   a. What is the distribution of $\frac{1}{100}\sum_{i=1}^{16}(X_i - 50)^2$?

   b. What is the distribution of $\bar{X}$?

6. Let $\bar{X}$ be the mean of a random sample of size 12 from the uniform distribution on the interval (0,1) (which has the mean 1/2 and variance 1/12). Approximate the probability $P(1/2 \leq \bar{X} \leq 2/3)$ using the central limit theorem.

7. Let the distribution of Y be $Binomial(25, 1/2)$. Find the probability $P(10 \leq Y \leq 12)$ in two ways: using the binomial pmf formula, and using the normal approximation. Comment on any difference.

8. The number X of flaws on a certain tape of length one yard follows a Poisson distribution with mean 0.3. We examine $n = 100$ such tapes and count the total number Y of flaws.

   a. Assuming independence, what is the distribution of Y?

   b. Find the exact and approximate probabilities for $P(Y \leq 25)$.

9. Let $Y = X_1 + X_2 + \cdots + X_{15}$ be the sum of a random sample of size 15 from the distribution whose pdf is $f(x) = (3/2)x^2$, $-1 < x < 1$. Approximate $P(-0.3 \leq Y \leq 1.5)$ using the central limit theorem.

10. Let the joint pmf of X and Y be defined by

$$f(x, y) = \frac{x + y}{32}, \quad x = 1, 2, \; y = 1, 2, 3, 4.$$

   a. Find the marginal pmf of X.

   b. Find the marginal pmf of Y.

   c. Calculate $P(X > Y)$.

   d. Calculate $P(Y = 2X)$.

   e. Calculate $P(X + Y = 3)$.

f. Calculate $P(X \leq 3 - Y)$.

g. Are X and Y independent?

h. Find $E(X)$.

i. Find $E(X + Y)$.

11. Two construction companies make bids of X and Y (in \$100,000's) on a remodeling project. The joint pmf of X and Y is uniform on the space $2 < x < 2.5$, $2 < y < 2.3$. If X and Y are within 0.1 of each other, the companies will be asked to rebid; otherwise the lower bidder will be awarded the contract. What is the probability that they will be asked to rebid?

12. Let $f(x, y) = 2e^{-x-y}$, $0 \leq x \leq y < \infty$, be the joint pdf of X and Y.

a. Find the marginal pdf $f_X(x)$ of X.

b. Find the marginal pdf $f_Y(y)$ of Y.

c. Compute $E(X)$ and $E(e^{-X-2Y})$.

d. Compute $P(X > \frac{1}{2})$.

e. Compute $P(X > \frac{1}{2}, Y > 2)$.

f. Compute $P(Y > 2|X > \frac{1}{2})$.

13. Let the joint pmf of X and Y be

$$f(x, y) = \frac{1}{4}, \quad (x, y) \in S = \{(0, 0), (1, 1), (1, -1), (2.0)\}.$$

a. Represent the joint pmf by a table.

b. Are X and Y independent?

c. Calculate $\text{Cov}(X, Y)$ and ρ.

14. Consider continuous random variables X and Y which have the following joint pdf

$$f(x, y) = 24xy, \quad x > 0, \ y > 0, \ x + y < 1.$$

a. Sketch a graph of the support of X and Y.

b. Find the probability $P(Y > 2X)$.

c. Find the marginal pdf $f_1(x)$ of X.

d. Find the mean $E(X)$.

e. Find the variance $\text{Var}(X)$.

f. Find the covariance $\text{Cov}(X, Y)$.

g. Find the correlation coefficient ρ between X and Y.

h. Find the conditional pdf $h(y|x)$ of Y given $X = x$.

i. Find the condition probability $P(Y \leq \frac{1}{3}(1-X)|X=x)$.

j. Find the conditional mean $E(Y|X=x)$.

15. Show that $\text{Cov}(aX+b, cX+d) = ac\text{Var}(X)$, where X is a random variable and a, b, c are deterministic constants.

16. Let the pmf of X be $f_1(x) = \frac{1}{10}$, $x = 0, 1, 2, \cdots, 9$, and the conditional pmf of Y given $X = x$ be $h(y|x) = \frac{1}{10-x}$, $y = x, x+1, \cdots, 9$. Find

    a. the joint pmf $f(x, y)$ of X and Y.

    b. The marginal pmf $f_2(y)$ of Y.

    c. $E(Y|x)$.

17. The marginal distribution of X is $U(0,1)$. The conditional distribution of Y, given $X = x$, is $U(0, e^x)$.

    a. Determine $h(y|x)$, the conditional pdf of Y, given $X = x$.

    b. Find $E(Y|x)$.

    c. Find the joint pdf of X and Y. Sketch the region where $f(x, y) > 0$.

    d. Find $f_2(y)$, the marginal pdf of Y.

    e. Find $g(x|y)$, the conditional pdf of X, given $Y = y$.

18. An obstetrician does ultrasound examinations on her patients between their 16th and 25th weeks of pregnancy to check on the growth of the unborn child. Let X equal the widest diameter of the head and Y be the length of the femur, both in mm. Assume that X and Y have a bivariate normal distribution with $\mu_X = 60.6$, $\sigma_X = 11.2$, $\mu_Y = 46.8$, $\sigma_Y = 8.4$, $\rho = 0.94$

    a. Find $P(40.5 < Y < 48.9)$

    b. Find $P(40.5 < Y < 48.9|X = 68.6)$

19. Karl Pearson carried out a famous study on the resemblances between fathers and sons. He found that the distribution from his sample of 1078 pairs of fathers and sons had a mean and standard deviation of height for fathers of 69 and 2 (in.), whereas for sons it was 70 and 2 (in.). The correlation between fathers' and sons' heights was 0.5. Assuming that the distribution can of the pairs of heights can be modelled as bivariate normal with the sample means, sds and correlation, what is

    a. the expected height for the son of a father who is 74 in. tall,

    b. the chance that the son's height is more than 1 in. from the expected height in (a)?

    c. the expected height of a father whose son is 72.5 in?

    d. the chance that both a father and son are above average height?

20. Scores in the two exams in a double weight course have means 65 and 60, standard deviations 18 and 20 and a correlation of 0.75. Assuming they can be modelled as bivariate normal, what is the expected score in the second exam for a student how is above average on the first exam?

21. Suppose (X, Y) has a standard bivariate normal density with correlation ρ. For a, b, find $E(Y|a < X < b)$.

22. Suppose (X, Y) are random variables and a, b, c, d are constants. Show

$$Cov(aX + bY, cX + dY) = acVar(X) + bdVar(Y) + (bc + da)Cov(X, Y).$$

23. Suppose that $X \sim N(0, 1)$ and that Z is an independent random variable which takes the values $1, -1$ each with probability $\frac{1}{2}$.

    a. Show that $X, -X$ have the same distribution.

    b. What is the distribution of ZX?

    c. Find $Cov(X, ZX), Corr(ZX, X)$.

    d. Are X and ZX independent? (Hint: consider $P(ZX \leq -1|X \leq -1)$.)

    e. Let $Y = ZX$. Is (X, Y) bivariate normal?

24. Suppose U has a uniform pdf on [0,1]. Find the pdf of $-\ln(1 - U)$.

25. Suppose X has the Laplace pdf, f, given by $f(x) = \frac{1}{2}e^{-|x|}$ for all real x.

    a. Find the cdf of X(hint: consider negative x first and compare with the exponential density).

    b. Find the cdf of X^2.

    c. Find the pdf of X^2.

    d. Find the cdf of X^4.

    e. Find the pdf of X^4.