MAST90105: Lab and Workshop Problems for Week 5

1. Suppose that 10^6 points are selected independently and at random from the unit square $\{(x, y) : 0 \le x \le 1, 0 \le y \le 1\}$. Let W equal the number of points that fall in $A = \{(x, y) : x^2 + y^2 < 1\}$.

   a. How is W distributed? Namely, what is the name of the distribution of W if it has a name? And what is the pmf of W?

   b. Give the mean, variance and standard deviation of W.

   c. What is the expected value of $W/250000$?

2. **The objective of this question is to use R to estimate the value of π using the idea and result of the previous question.** (*Note that everybody knows the meaning of π, but nobody can write down the exact and explicit decimal representation of π. So estimating the value of π is meaningful even though the method used here is clearly not the best.*)

   Note that selecting at random a million values of x from [0,1] can be done using `runif(`10^6`)` command in R.

   a. Using this information and the knowledge that you have learned about R so far, try to write up a few lines of R commands to generate an observation of W. Remember that your R commands must not involve the use of the true value of π.

      (Note that a command like `mean(rbinom(1000,`10^6`,pi/4))` can be used to estimate π. But this is not the one we want because it involves the use of the true value of π. Use Help inside R)

   b. Once you get an observation of W, calculate $W/250000$ and see how the result is close to π.

   c. Think about how you can improve the precision of your estimate of π.

      One way of doing this would be to implement the R commands developed by you into an R function. Then use this function to generate a number of W observations. Then use the average of the generated W values divided by 250000 to estimate π.

3. **The objective of this question is to use R to produce plots of Gamma densities, introducing to the Graphics package, ggplot2, which is a standard for professional graphics.**

   The idea of ggplot2 is to start with a base plot and add the other elements of your plot incrementally. A quick guide is in this link. The code is available in Lab5.R in the LMS Labs folder.

   Here is the code used to produce one of the plots in Lectures:

```
p <- ggplot(data = data.frame(x= c(0,30)),aes(x=x)) +
stat_function(fun=function(x)dgamma(shape=0.25,x,scale=4),aes(colour = "1"))
```

```
+
stat_function(fun = function(x)dgamma(shape=1,x,scale=4), aes(colour = "2"))
+
stat_function(fun = function(x)dgamma(shape=2,x,scale=4), aes(colour = "3"))
*
stat_function(fun = function(x)dgamma(shape=3,x,scale=4), aes(colour = "4"))
+
newcols <- c("1"="red","2"="blue","3"="darkgreen","4"="purple")
p +scale_colour_manual(values = newcols,name = "",labels = c(expression("
" * alpha==0.25*" "),
expression(" "*alpha==1*" " ),
expression(" "*alpha==2*" " ),
expression(" "*alpha==3*" " ))) +
theme(legend.position="top",
text=element_text(size=22),
panel.background =element_rect(fill="white"),
axis.line = element_line(colour = "black") ) +
ylim(0,0.25)+
ylab("f(x)") +
annotate("text",x=15,y=0.15, label="theta==4",parse=TRUE, size=8)+
ggtitle("Gamma Probability Density Functions - Varying Shape")
```

a. Find out about the definition of data.frame and dgamma. How are these used in the code?

b. Alter the code so it produces Gamma densities with shape parameters 0.1, 1.1, 5 and 10 and scale = 5. You will have to alter a number of parts of the code. Comment on the results in comparison to those in lectures.

c. Alter the code so it produces Gamma densities with scale parameters 0.1, 1.1, 5 and 10 and shape = 5. You will have to alter a number of parts of the code. Comment on the results in comparison to those in lectures.

d. Find out how the legends are plotted and write down which parts of the code are involved in the production of the legends.

4. Let the random variable X have the pdf $f(x) = 2(1-x)$, $0 \le x \le 1$, 0 elsewhere.

   a. Sketch the graph of this pdf.

   b. Determine and sketch the graph of the distribution function of X.

   c. Find

      i. $P(0 \le X \le 1/2)$,

      ii. $P(1/4 \le X \le 3/4)$,

      iii. $P(1/4 \le X \le 5/4)$,

      iv. $P(X = 3/4)$,

      v. $P(X \ge 3/4)$,

vi. the value of μ,
vii. the value of σ^2, and
viii. the 36th percentile $\pi_{0.36}$ of X.

5. The pdf of X is $f(x) = c/x^2$, $1 < x < \infty$.
   a. Find the value of c so that $f(x)$ is a pdf.
   b. Show that $E(X)$ is not finite.

6. Let $f(x) = 1/2$, $0 < x < 1$ or $2 < x < 3$, 0 elsewhere, be the pdf of X.
   a. Sketch the graph of this pdf.
   b. Define cdf of X and sketch its graph.
   c. Find $q_1 = \pi_{0.25}$.
   d. Find $m = \pi_{0.50}$. Is it unique?
   e. Find the value of $E(X)$.

7. Let $F(x) = 1 - (\frac{1}{2}x^2 + x + 1)e^{-x}$, $0 < x < \infty$ be the cdf of X.
   a. Find the mgf $M(t)$ of X.
   b. Find the values of μ and σ^2.

8. The life X in years of a voltage regulator of a car has the pdf

$$f(x) = \frac{3x^2}{7^3}e^{-(x/7)^3}, \quad 0 < x < \infty.$$

   a. What is the probability that this regulator will last at least 7 years?
   b. Given that it has lasted at least 7 years, what is the conditional probability it will last at least another 3.5 years?

   A function is given as $f(x) = (3/16)x^2$, $-c < x < c$.
   a. Find the constant c so that $f(x)$ is a pdf of a random variable X.
   b. Find the cdf $F(x) = P(X \leq x)$.
   c. Sketch graphs of the pdf $f(x)$ and the cdf $F(x)$.

9. A function is given as $f(x) = 4x^c$, $0 \leq x \leq 1$.
   a. Find the constant c so that $f(x)$ is a pdf of a random variable X.
   b. Find the cdf $F(x) = P(X \leq x)$.
   c. Sketch graphs of the pdf $f(x)$ and the cdf $F(x)$.

10. Customers arrive randomly at a bank teller's window. Given that one customer arrived during a particular 10-minute period, let X equal the time within the 10 minute-period that the customer arrived. Assuming X is $U(0, 10)$, find:

a. the pdf of X,

b. $P(X \geq 8)$,

c. $P(2 \leq X < 8)$,

d. $E(X)$, and

e. $\text{Var}(X)$.

11. Let X have an exponential distribution with a mean of $\theta = 20$. Compute

    a. $P(10 < X < 30)$,

    b. $P(X > 30)$.

    c. $P(X > 40 | X > 10)$.

    d. What are the variance and the mgf of X?

    e. Find the 80th percentile of X.

12. What are the pdf, the mean, and the variance of X if the mgf of X is given by the following?

    a. $M(t) = (1 - 3t)^{-1},\ t < 1/3$.

    b. $M(t) = \frac{3}{3-t},\ t < 3$.

13. Let X_t equal the number of flawed recordings in each length t measured in billions of records. Assume that X_t is a Poisson process with rate 2.5 per billion records (so $t$is treated as continuous). Let W be the length of records before the first bad record is found.

    a. Give the mean number of flaws per billion records.

    b. How is W distributed?

    c. Give the mean and variance of W.

14. Let random variable X have the pdf

$$f(x) = \frac{e^{-x}}{(1 + e^{-x})^2}, \quad -\infty < x < \infty.$$

(The distribution of such X is known as the *logistic distribution*.)

a. Write down the cdf of X.

b. Find the mean and variance of X.

c. Find $P(3 < X < 5)$.

d. Find the 85-th percentile of X.

e. Let $Y = \frac{1}{1+e^{-X}}$. Find the cdf of Y.
Can you tell the name of the distribution of Y?

15. Telephone calls enter a university switchboard at a mean rate of 2/3 call per minute according to a Poisson process. Let X denote the waiting time until the 10th call arrives.

    a. What is the pdf of X?.

    b. What are the mgf, mean and variance of X?

16. If X has a gamma distribution with scale parameter $\theta = 4$ and $\alpha = 2$, find $P(X < 5)$.

17. Use an argument similar to the one used for the exponential distribution, find the moment generating function of the gamma distribution with general non-integer scale parameter. Use the moment-generating function of a gamma distribution to show that $E(X) = \alpha\theta$ and $\mathrm{Var}(X) = \alpha\theta^2$.

18. Let X have a $\chi^2(2)$ distribution. Find constants a and b such that

$$P(a < X < b) = 0.90, \quad \text{and} \quad P(X < a) = 0.05.$$