ECO239 Statistics I R Homework 2

DUE: December 24th (Thursday) at 16:00
Submit a printed hard-copy

Just staple them, don't place inside a plastic cover, to save time and resource. Thanks!

- The objective of this task is for you to be familiar with distributions we have studied in Ch.4 and Ch.5.
- The tasks can be done either by using R console window or by installing "RStudio".
- For reporting, copy & paste figures, codes & output, in addition to your comments.

*RStudio can be downloaded from https://www.rstudio.com/products/RStudio/

Basics:

- **Codes for each distribution**:

Distribution		Distribution	
Binomial	*binom*	Uniform	*unif*
Hypergeometric	*hyper*	Normal	*norm*
Poisson	*pois*	Exponential	*exp*

| "d"	returns the height of the probability density function
"p"	returns the cumulative density function

e.g. *dbinom* returns the probability density function for Binomial distribution.
e.g. *pnorm* returns the cumulative density function for Normal distribution.

- **Finding cumulative probabilities**: x_0 is the value of your interest. You should fill in blanks to derive actual probabilities.

Binomial	*pbinom(c(x_0), size = , prob= , lower.tail= [TRUE/FAULSE])*
Hypergeometric	*phyper(c(x_0), m= , n= , k= , lower.tail=[TRUE/FAULSE])*
Poisson	*ppois(c(x_0), lambda= , lower.tail=[TRUE/FAULSE])*
Uniform	*punif(c(x_0),min= , max= , lower.tail=[TRUE/FAULSE])*
Normal	*pnorm(c(x_0),mean= , sd= ,lower.tail=[TRUE/FAULSE])*
Exponential	*pexp(c(x_0),rate= , lower.tail=[TRUE/FAULSE])*

Note: *lower.tail=TRUE* if you want to derive the lower-tail probability.

lower tail   upper tail
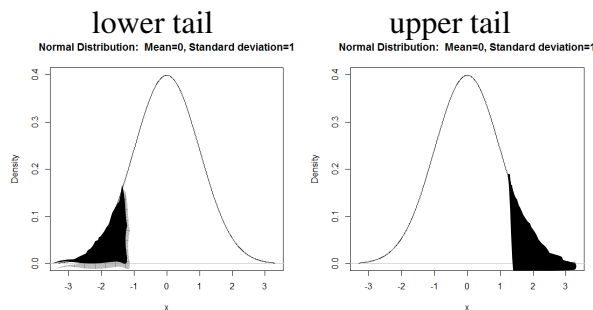
lower.tail=FAULSE if you want to derive the upper-tail probability.

- **Plots**

Basic form:

define x values with its minimum, maximum and increment. Set the increment as 0.01.

x<- seq(min, max, increment)

plot(x, dbinom(x, size=n, prob=p),type="h", xlab="Number of Successes", ylab="Probability Mass",main="Binomial Distribution:Trials=n, Probability of success=p")

*Change the entries inside the parenthesis after d_____ (n, p, name of the distribution), labels for x and y axes and main title according to the problem you are working on.

Task1: Discrete Distributions and Probabilities

I: Binomial Distribution

a. Plot probability mass function for Binomial Distribution with the probability of success p = 0.4 and n = 5, 25, 50. When you change p, how does it affect the shapes of the distribution? Comment.
   [Hint: *set x<-seq(0,40,0.01)for all plots*]

b. Create a table for the marginal Binomial probabilities when p = 0.4 and n = 10.
   [Hint: *dbinom(0:n, size=n, prob=p)]*

c. Create a table for the cumulative Binomial probabilities when p = 0.4 and n = 10.
   [Hint: *pbinom(0:n, size=n,prob=p)]*

d. 25% of students admitted to University A will actually enroll.
   i. What's the probability that less than 8 students will enroll if 25 students were offered admissions?
   ii. What's the probability that more than 9 students will enroll if 25 students were offered admissions?
   iii. If 60% of students admitted to University will actually enroll, what will be the results for i and ii?
   [Hint: Use cumulative distribution. Pay attention to "less than" and "more than" expressions!]

II: Hypergeometric Distribution

- m = the number of white balls in the urn (= S)
- n = the number of black balls in the urn (= N-S)
- k = the number of balls drawn from the urn (= n)

e. Plot probability mass function for Hypergeometric Distribution with N=100, n = 10, S = 10, 30 and 50. When you change the value of S, how does it affect the shapes of the distribution? Comment.
   [Hint: *set x<-seq(0,10,0.01) for all plots*]

f. A company receives a shipment of 100 items. A random sample of 10 items is selected and the shipment is rejected if at least one of these item is defective.

      i.   What is the probability of finding at most 2 defectives if the shipment contains 12 defectives?

      ii.   What is the probability of rejecting a shipment containing 15 defectives?

III: Poisson Distribution

g. Plot probability mass function for Poisson Distribution with the expected number of success per time unit = 0.5, 2, 10. When you change the value of lambda, how does it affect the shapes of the distribution? Comment.
[Hint: set x<- seq(0,22,0.01) for all plots.]

h. The number of accidents in a production facility has a Poisson distribution with a mean of 1.8 per month.

      i.   For a given month what is the probability there will be less than 2 accidents?

      ii.   What is the probability there will be more than 15 accidents <u>per year</u>?

i. An insurance company reported that 5.4% of claims forms submitted for payment after a complex surgical procedure contain errors. If 100 of these forms are chosen at random, what is the probability that fewer than 3 of them contain errors?

      i.   Use Binomial Distribution to find the exact answer.

      ii.   Use Poisson approximation to find an approximate answer.

      iii.   Is Poission Approximation working well? Explain.

Task2: Continuous Distribution and Probabilities

IV: Uniform Distribution

j. Plot probability density functions (pdf) for Uniform distribution in the range of [0 10]. Change the range and observe the changes. Comment.
[Hint: set x<-seq(min, max, 0.01) by changing min=0,max=10 for this question. Use type="l" instead of type="h" for continuous distributions]

k. Calculate the probability $P(x > 2)$ when x~U(0,10).

l. The jurisdiction of a rescue team includes emergencies occurring on a stretch of river that is 5 miles long. Experience has shown that the distance along this stretch, measured in miles from its northernmost point, at which an emergency occurs can be represented by a uniformly distributed random variable over the range 0 to 5 miles. Then, if X denotes the distance of an emergency from the northernmost point of this stretch of river, its probability density function is

$$f(x) = \begin{Bmatrix} 0.2 \text{ for } 0 < x < 5 \\ 0 \text{ for all other x} \end{Bmatrix}$$

      i.   Find the probability that a given emergency arises within 1 mile of the northernmost point of this stretch of river.

      ii.   The rescue team's base is at the midpoint of this stretch of river. Find the probability that a given emergency arises more than 1.5 miles from this base.

V: Normal Distribution

m. Plot pdfs for Normal distribution when population mean is 0 and the standard deviation is 1 (= standard normal distribution). Change the value of standard deviation to 10 and observe the changes in the shapes of the distribution. Change the value of mean to 10 (the standard deviation =10) and observe the changes in the shapes of the distribution. Comment.
[Hint: set x<- seq(-30,30,0.01) for all plots.Use type="l" for the plot.]

n. Calculate the probability $P(x < 35)$ when x~ N(60, 121).

o. Calculate the probability $P(x > 72)$ when x ~ N(60,121).

p. It is estimated that the time that a well-known pianist, Fazıl Say, spends on stage at his concerts follows a normal distribution with a mean of 100 minutes and a standard deviation of 15 minutes.

    i.   What is the probability that his concerts lasts between 110 and 130 minutes?

    ii.   An audience is attempting to record the entire concert using her cell phone, which has a capacity of recording 120 minutes. What is the probability that this capacity will be insufficient to record the entire concert?

    iii.   If the standard deviation of concert time was only 10 minutes, what is the probability that a concert would last more than 120 minutes? Compare your answer with the one from (ii).

q. A car rental company has determined that the probability a car will need service work in any given month is 0.2. The company has 500 cars. What is the probability that fewer than 90 cars will require service work in a particular month?

    i.   Solve by using binomial distribution.

    ii.   Solve by using normal approximation

    iii.   Is normal approximation a good approximation? Explain.

VI: Exponential Distribution

r. Plot pdfs for Exponential distribution when the lambda parameter is 0.2. Change the parameter to 0.5 and observe the changes in the shapes. Comment.
[Hint: set x<-seq(0,30,0.01) for all plots.]

s. Calculate the probability $P(t < 2)$ when t~ Exp(lambda = 0.3).

t. Calculate the probability $P(t > 3)$ when t~ Exp(lambda = 0.3).

u. Delivery trucks arrive independently at PTT distribution center. The mean number of trucks arriving per hour is 30. Given that the truck has just arrived, answer the following questions.

    i.   What is the probability that the next truck will arrive within the next 3 minute?

    ii.   What is the probability that the next truck will arrive between 6 and 12 minute?