604_HW2

Daina Bouquin

```
library(ggplot2)
library(dplyr)
```

1. Suppose that X is a discrete random variable having probability function $Pr(X = k) = ck^2$ for $k = 1, 2, 3$. Find $c, Pr(X \leq 2), E[X]$, and $Var(X)$.

$$Pr(X = 1) + Pr(X = 2) + Pr(X = 3) = 1$$

$$1^2c + 2^2c + 3^2c = 1$$

$$1c + 4c + 9c = 1$$

$$14c = 1$$

$$c = 1/14$$

```
# c = 1/14 therefore:
PrX1 <- (1/14)*(1^2)
PrX1
```

```
## [1] 0.07142857
```

```
PrX2 <- (1/14)*(2^2)
PrX2
```

```
## [1] 0.2857143
```

```
PrX3 <- (1/14)*(3^2)
PrX3
```

```
## [1] 0.6428571
```

```
# Pr(X ≤ 2)
PrX1 + PrX2
```

```
## [1] 0.3571429
```

```
#Expected value
E_X <- PrX1 + (PrX2 * 2) + (PrX3 * 3)
E_X
```

```
## [1] 2.571429
```

```
# Var(X) = E[X^2] - (E[X])^2
var1 <- PrX1 + (PrX2 * 2^2) + (PrX3 * 3^2)
var2 <- (PrX1 + (PrX2 * 2) + (PrX3 * 3))^2
var1 - var2
```

```
## [1] 0.3877551
```

2. Suppose that X is a continuous random variable having p.d.f. $f(x) = cx^2$ for $1 \leq x \leq 2$. Find $c, Pr(X \geq 1), E[X]$, and $Var(X)$.

$$\int_1^2 cx^2 = 1 = \frac{1}{3}cx^3$$

$$\frac{1}{3}c(2)^3 - \frac{1c(1)^3}{3} = 1$$

$$\frac{8c}{3} - \frac{c}{3}$$

$$c = \frac{3}{7}$$

Given this range, $Pr(x \geq 1) = 1$

```
# Expected value
# f(x) = cx^2
# E(X) = Integral of xf(x)dx over the whole domain of X.

f <- function(x){
     (3/7)*x^2
}

f2 <- function(x){x*f(x)}

integrate(f2,1,2)
```

```
## 1.607143 with absolute error < 1.8e-14
```

```
# Var(X) = E[X^2] - (E[X])^2
f <- function(x){
     (3/7)*x^4
}
f3 <- function(x){x*f(x)}

integrate(f3,1,2)
```

```
## 4.5 with absolute error < 5e-14
```

```
(4.5) - ((1.6)^2)
```

```
## [1] 1.94
```

3. Suppose that X and Y are jointly continuous random variables with:

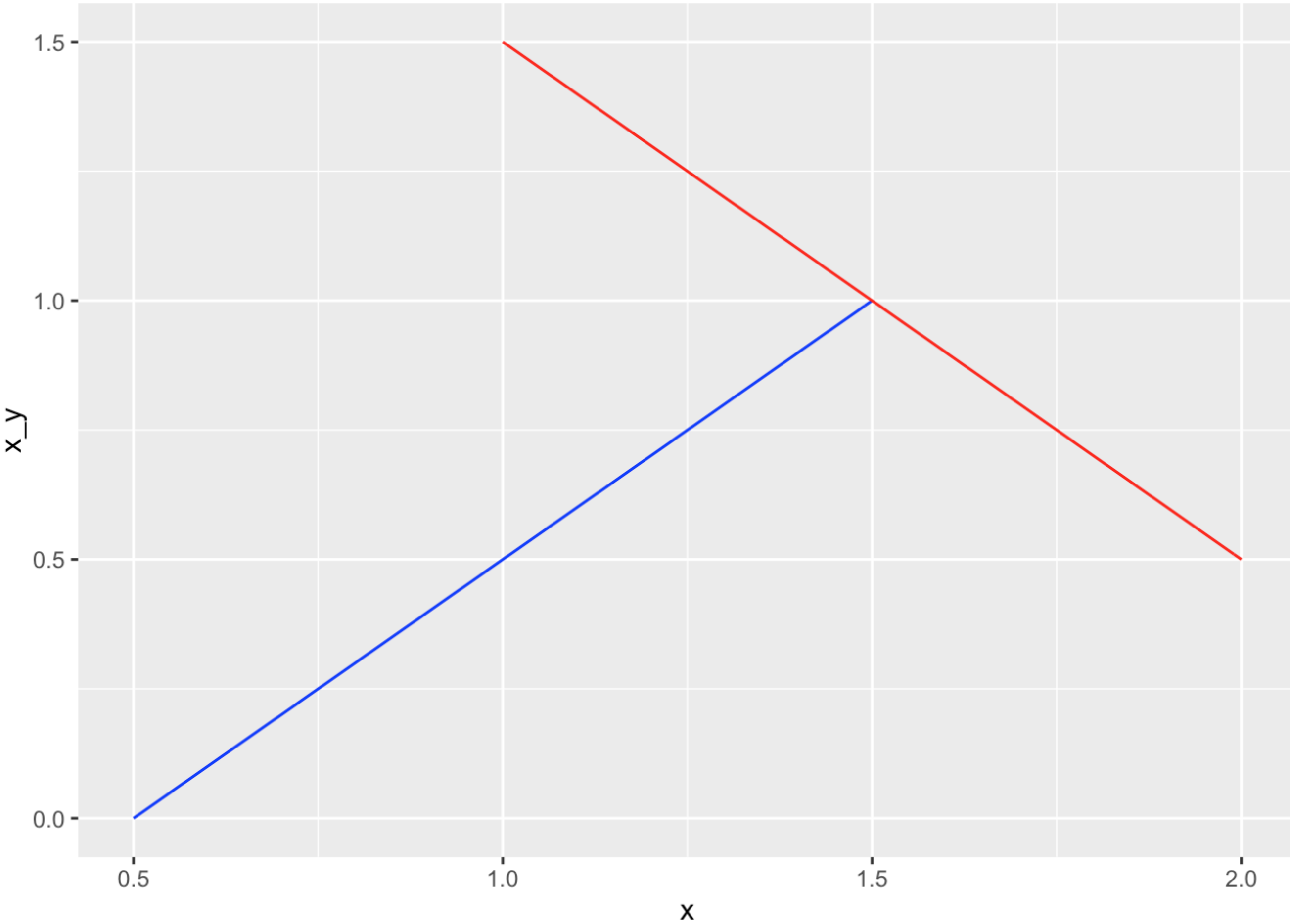
$$\begin{cases} y - x & \text{for } 0 < x < 1 \text{ and } 1 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

a. Compute and plot $f_x(x)$ and $f_y(y)$

```
x_y <- seq(from = 0, to = 1, by = 0.01)
y_x <- seq(from = 1, to = 2, by = 0.01)
x <- y_x - (1/2)
y <- (3/2) - x_y

# create a dataframe for the variables
J_randvar <- as.data.frame(cbind(x_y, x, y_x, y))

ggplot(J_randvar) + geom_line(aes(x=x, y=x_y), colour="blue") + geom_line(aes(x=y_x, y=y),
 colour = "red")
```



b. Are X and Y independent?
   X and Y are not independent.
c. Compute $F_X(x)$ and $F_Y(y)$

$$F_X(x) = \int f_X(x)dx = \int \frac{1}{2}y^2 - xydx = \int \frac{1}{2}y^2\,dx - \int xydx = \frac{1}{2}xy^2 - \frac{x^2y}{2}$$

$$F_Y(y) = \int f_Y(y)dy = \int xy - \frac{1}{2}x^2\,dy = \int xydy - \int \frac{1}{2}x^2\,dy = \frac{xy^2}{2} - \frac{1}{2}yx^2$$

d. Compute $E[X], Var(X), E[Y], Var(Y), Cov(X, Y)$, and $Corr(X, Y)$.
   …

4. Suppose that the following 10 observations come from some distribution (not highly skewed) with unknown mean μ.
   7.3, 6.1, 3.8, 8.4, 6.9, 7.1, 5.3, 8.2, 4.9, 5.8
   Compute $\bar{X}, S2$, and an approximate 95% confidence interval for μ.

```
obs <- c(7.3, 6.1, 3.8, 8.4, 6.9, 7.1, 5.3, 8.2, 4.9, 5.8)
(X_ <- mean(obs))
```

```
## [1] 6.38
```

```
(S2 <- cov(obs, obs)) #?cov = compute the variance of x and the covariance or correlatio
n of x and y
```

```
## [1] 2.161778
```

```
t.test(obs) # 95 percent confidence interval: 5.328212 7.431788
```

```
##
##  One Sample t-test
##
## data:  obs
## t = 13.722, df = 9, p-value = 2.439e-07
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  5.328212 7.431788
## sample estimates:
## mean of x
##      6.38
```

5. A random variable X has the memoryless property if, for all $s, t > 0$, $Pr(X > t + s|X > t) = Pr(X > s)$
Show that the exponential distribution has the memoryless property.
pg. 193 DES - memoryless property:
for all s ≥ 0 and t ≥ 0:

$$P(X > s + t|X > s) = \frac{P(X > s + t)}{P(X > s)}$$

$$P(X > s + t|X > s) = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t}$$

$$= P(X > t)$$

6. Suppose $X1, X2, \ldots, Xn$ are i.i.d. $Exp(\lambda = 1)$. Use the Central Limit Theorem to find the approximate value of $Pr(100 \leq \sum 100X \leq 110).\, i = 1i$

```
# CLT - given certain conditions, the mean of a sufficiently large number of iterates of
 independent random variables, each with a well-defined (finite) expected value and fini
te variance, will be approximately normally distributed, regardless of the underlying di
stribution.
# Show prob of a normal distribution with mean= 100 and var= 100 between 100 and 110.
set.seed(250)
prob6_110 <- pnorm(110, mean = 100, sd = 10, lower.tail = TRUE)
prob6_100 <- pnorm(100, mean = 100, sd = 10, lower.tail = TRUE)
(prob6_110 - prob6_100)
```

```
## [1] 0.3413447
```

5.13

A random variable X that has a pmf given by $p(x) = 1/(n+1)$ over the range $R_X = 0, 1, 2, \ldots n$ is said to have a discrete uniform distribution.

a. Find the mean and variance of this distribution

$$E[X] = (\frac{n(n+1)}{2})(\frac{1}{(n+1)}) = \frac{n}{2}$$

$$Var[x] = E[X^2] - E[X]^2$$

$$Var[x] = (\frac{n(n+1)(2n+1)}{6})(\frac{1}{(n+1)}) - (\frac{n}{2})^2$$

$$\ldots$$

$$= \frac{n(n+2)}{12}$$

b. If $R_X = (a, a+1, a+2, \ldots, b)$, compute the mean and variance of X.

…

5.14

integral of pdf = $-e^{(-0.4x)}$

```
sat_pdf <- function(years){return(-exp(-0.4 * years))}
# 5 year prob
(P_yr.5 <-  1 - (sat_pdf(5) - sat_pdf(0)))
```

```
## [1] 0.1353353
```

```
# 3-6 year prob
(P_yr.3.6 <- sat_pdf(6) - sat_pdf(3))
```

```
## [1] 0.2104763
```

5.39

a. What is the distribution of the length of the linkage?

```
link_len <- function(){
  l <- rnorm(1, 60, sqrt(0.09)) + rnorm(1, 40, sqrt(0.05)) + rnorm(1, 50, sqrt(0.11))
  return(l)
}
# simulate 3000
set.seed(200)
sim_link_len <- replicate(3000, link_len())
# summarize the simulation results
summary(sim_link_len)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   148.3   149.7   150.0   150.0   150.3   151.8
```

```
# show variance
var(sim_link_len)
```

```
## [1] 0.2515157
```

b. What is the probability that the linkage will be longer than 150.2 cm?

```
set.seed(200)
(1 - pnorm(150.2, mean = 150, sd = sqrt(0.25)))
```

```
## [1] 0.3445783
```

c. The tolerance limites for the assembly are (149.83, 150.21). What proportion of assemblies are within the tolerance limits?

```
set.seed(200)
pnorm(150.21, mean = 150, sd = sqrt(0.25)) - pnorm(149.83, mean = 150, sd = sqrt(0.25))
```

```
## [1] 0.295829
```