

Multivariate Expectation Application Solutions

Lt Col Ken Horton

Lt Col Kris Pruitt

Professor Bradley Warner

20 October, 2020

Exercises

1. Let X and Y be continuous random variables with joint pdf:

$$f_{X,Y}(x,y) = x + y$$

where $0 \leq x \leq 1$ and $0 \leq y \leq 1$.

- a. Find $E(X)$ and $E(Y)$. We will use the marginal pdfs found in the Application 14 solution.

$$E(X) = \int_0^1 x \left(x + \frac{1}{2} \right) dx = \frac{x^3}{3} + \frac{x^2}{4} \Big|_0^1 = \frac{1}{3} + \frac{1}{4} = \frac{7}{12} = 0.583$$

Or numerically:

```
f <- function(x) { x[1]*(x[1] + x[2]) } # "x" is vector
adaptIntegrate(f, lowerLimit = c(0, 0), upperLimit = c(1, 1))
```

```
## $integral
## [1] 0.5833333
##
## $error
## [1] 1.110223e-16
##
## $functionEvaluations
## [1] 17
##
## $returnCode
## [1] 0
```

$$E(Y) = \int_0^1 y \left(y + \frac{1}{2} \right) dy = 0.583$$

- b. Find $\text{Var}(X)$ and $\text{Var}(Y)$.

$$\begin{aligned} \text{Var}(X) &= E(X^2) - E(X)^2 \\ E(X^2) &= \int_0^1 x^2 \left(x + \frac{1}{2} \right) dx = \frac{x^4}{4} + \frac{x^3}{6} \Big|_0^1 = \frac{1}{4} + \frac{1}{6} = \frac{5}{12} = 0.417 \end{aligned}$$

```
f <- function(x) { x[1]^2*(x[1] + x[2]) } # "x" is vector
round(adaptIntegrate(f, lowerLimit = c(0, 0), upperLimit = c(1, 1))$integral,3)
```

```
## [1] 0.417
```

So, $\text{Var}(X) = 0.417 - 0.583^2 = 0.076$.

Similarly, $\text{Var}(Y) = 0.076$.

c. Find $\text{Cov}(X, Y)$ and ρ . Are X and Y independent?

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

$$\begin{aligned} E(XY) &= \int_0^1 \int_0^1 xy(x+y) dy dx = \int_0^1 \left. \frac{x^2 y^2}{2} + \frac{xy^3}{3} \right|_0^1 dx = \int_0^1 \frac{x^2}{2} + \frac{x}{3} dx \\ &= \left. \frac{x^3}{6} + \frac{x^2}{6} \right|_0^1 = \frac{1}{3} = 0.333 \end{aligned}$$

```
f <- function(x) { x[1]*x[2]*(x[1] + x[2]) } # "x" is vector
round(adaptIntegrate(f, lowerLimit = c(0, 0), upperLimit = c(1, 1))$integral,3)
```

```
## [1] 0.333
```

So,

$$\text{Cov}(X, Y) = \frac{1}{3} - \left(\frac{7}{12} \right)^2 = -0.007$$

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{-0.007}{\sqrt{0.076 \times 0.076}} = -0.909$$

With a non-zero covariance, X and Y are not independent.

d. Find $\text{Var}(3X + 2Y)$.

$$\begin{aligned} \text{Var}(3X + 2Y) &= \text{Var}(3X) + \text{Var}(2Y) + 2\text{Cov}(3X, 2Y) = 9\text{Var}(X) + 4\text{Var}(Y) + 12\text{Cov}(X, Y) \\ &= 9 * 0.076 + 4 * 0.076 + 12 * -0.007 = 0.910 \end{aligned}$$

2. Optional - not difficult but does have small Calc III idea. Let X and Y be continuous random variables with joint pmf:

$$f_{X,Y}(x, y) = 1$$

where $0 \leq x \leq 1$ and $0 \leq y \leq 2x$.

a. Find $E(X)$ and $E(Y)$.

$$E(X) = \int_0^1 x \cdot 2x dx = \left. \frac{2x^3}{3} \right|_0^1 = 0.667$$

$$E(Y) = \int_0^2 y \left(1 - \frac{y}{2}\right) dy = \frac{y^2}{2} - \frac{y^3}{6} \Big|_0^2 = 2 - \frac{8}{6} = 0.667$$

b. Find $\text{Var}(X)$ and $\text{Var}(Y)$.

$$E(X^2) = \int_0^1 x^2 \cdot 2x dx = \frac{x^4}{2} \Big|_0^1 = 0.5$$

$$\text{So, } \text{Var}(X) = 0.5 - \left(\frac{2}{3}\right)^2 = \frac{1}{18} = 0.056$$

$$E(Y^2) = \int_0^2 y^2 \left(1 - \frac{y}{2}\right) dy = \frac{y^3}{3} - \frac{y^4}{8} \Big|_0^2 = \frac{8}{3} - 2 = 0.667$$

$$\text{So, } \text{Var}(Y) = \frac{2}{3} - \left(\frac{2}{3}\right)^2 = \frac{2}{9} = 0.222$$

c. Find $\text{Cov}(X, Y)$ and ρ . Are X and Y independent?

$$E(XY) = \int_0^1 \int_0^{2x} xy dy dx = \int_0^1 \frac{xy^2}{2} \Big|_0^{2x} dx = \int_0^1 2x^3 dx = \frac{x^4}{2} \Big|_0^1 = \frac{1}{2}$$

So,

$$\text{Cov}(X, Y) = \frac{1}{2} - \frac{2}{3} \frac{2}{3} = \frac{1}{18} = 0.056$$

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{\frac{1}{18}}{\sqrt{\frac{1}{18} \frac{2}{9}}} = 0.5$$

X and Y appear to be positively correlated (thus not independent).

d. Find $\text{Var}\left(\frac{X}{2} + 2Y\right)$.

$$\text{Var}\left(\frac{X}{2} + 2Y\right) = \frac{1}{4}\text{Var}(X) + 4\text{Var}(Y) + 2\text{Cov}(X, Y) = \frac{1}{72} + \frac{8}{9} + \frac{1}{9} = 1.014$$

3. Suppose X and Y are *independent* random variables. Show that $E(XY) = E(X)E(Y)$.

If X and Y are independent, then $\text{Cov}(X, Y) = 0$. So,

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$$

Thus,

$$E(XY) = E(X)E(Y)$$

4. You are playing a game with a friend. Each of you roll a fair sided die and record the result.

a. Write the joint probability mass function.

Let X be the number on your die and Y be the number on your friend's die.

		X					
		1	2	3	4	5	6
Y	1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
	2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
	3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
	4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
	5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
	6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

b. Find the expected value of the product of your score and your friend's score.

To find $E[XY]$, we determine all 36 values of the product of X and Y and multiply by the associated probabilities. Since the probabilities are all equal, we will take the $\frac{1}{36}$ out of the summation. Now

$$\begin{aligned}
 E[XY] &= \frac{1}{36}(1 + 2 + 3 + 4 + 5 + 6 + 2 + 4 + \\
 &\quad 6 + 8 + 10 + 12 + 3 + 6 + 9 + 12 + 15 + 18 + 4 + 8 + 12 + 16 + 20 + 24 + \\
 &\quad 5 + 10 + 15 + 20 + 25 + 30 + 6 + 12 + 18 + 24 + 30 + 36) \\
 &= 12.25
 \end{aligned}$$

c. Verify the previous part using simulation.

```
set.seed(1012)
(do(100000)*(sample(1:6,size=2,replace=TRUE))) %>%
  mutate(prod=V1*V2) %>%
  summarize(Expec=mean(prod))
```

```
##      Expec
## 1 12.25016
```

d. Using simulation, find the expected value of the maximum number on the two rolls.

```
(do(100000)*max(sample(1:6,size=2,replace=TRUE))) %>%
  summarize(Expec=mean(max))
```

```
##      Expec
## 1 4.4737
```

5. A miner is trapped in a mine containing three doors. The first door leads to a tunnel that takes him to safety after two hours of travel. The second door leads to a tunnel that returns him to the mine after three hours of travel. The third door leads to a tunnel that returns him to his mine after five hours. Assuming that the miner is at all times equally likely to choose any one of the doors, yes a bad assumption but it makes for a nice problem, what is the expected length of time until the miner reaches safety?

Simulating this is a little more challenging because we need a conditional but we try it first before going to the mathematical solution.

Let's write a function that takes a vector and returns the sum of the values up to the first time the number 2 appears, we are using the time values as our sample space. Anytime you are repeating something more than 5 times, it might make sense to write a function.

```
miner_time <- function(x){
  index <- which(x==2)[1]
  total<-cumsum(x)
  return(total[index])
}

set.seed(113)
(do(10000)*miner_time(sample(c(2,3,5),size=20,replace=TRUE))) %>%
  summarise(Exp=mean(miner_time))

##      Exp
## 1 10.0092
```

Now let's find it mathematically.

Let X be the time it takes and Y the door. Then we have

$$\begin{aligned} E[X] &= E[E[X|Y]] \\ &= \frac{1}{3}E[X|Y=1] + \frac{1}{3}E[X|Y=2] + \frac{1}{3}E[X|Y=3] \end{aligned}$$

Now if door 2 is selected

$$E[X|Y=2] = E[X] + 3$$

since the miner will travel for 3 hours and then be back at the starting point.

Likewise if door 3 is select

$$E[X|Y=3] = E[X] + 5$$

So

$$E[x] = \frac{1}{3}2 + \frac{1}{3}(E[X] + 3) + \frac{1}{3}(E[X] + 5)$$

$$\begin{aligned} E[x] - \frac{2}{3}E[X] &= \frac{2}{3} + \frac{3}{3} + \frac{5}{3} \\ \frac{1}{3}E[X] &= \frac{10}{3} \end{aligned}$$

$$E[X] = 10$$

6. ADVANCED: Let X_1, X_2, \dots, X_n be independent, identically distributed random variables. (This is often abbreviated as “iid”). Each X_i has mean μ and variance σ^2 (i.e., for all i , $E(X_i) = \mu$ and $\text{Var}(X_i) = \sigma^2$).

Let $S = X_1 + X_2 + \dots + X_n = \sum_{i=1}^n X_i$. And let $\bar{X} = \sum_{i=1}^n \frac{X_i}{n}$.

Find $E(S)$, $\text{Var}(S)$, $E(\bar{X})$ and $\text{Var}(\bar{X})$.

$$E(S) = E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n) = \mu + \mu + \dots + \mu = n\mu$$

Since the X_i s are all independent:

$$\text{Var}(S) = \text{Var}(X_1 + X_2 + \dots + X_n) = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n) = n\sigma^2$$

$$E(\bar{X}) = \frac{1}{n}E(X_1 + X_2 + \dots + X_n) = \frac{1}{n}n\mu = \mu$$

$$\text{Var}(\bar{X}) = \frac{1}{n^2}\text{Var}(X_1 + X_2 + \dots + X_n) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

File Creation Information

- File creation date: 2020-10-20
- Windows version: Windows 10 x64 (build 18362)
- R version 3.6.3 (2020-02-29)
- `mosaic` package version: 1.7.0
- `tidyverse` package version: 1.3.0