Multivariate Expectation Application Solutions

Lt Col Ken Horton   Lt Col Kris Pruitt   Professor Bradley Warner

21 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 $\mathrm{E}(X)$ and $\mathrm{E}(Y)$. We will use the marginal pdfs found in the Application 14 solution.

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

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
```

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

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

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$$

$$\mathrm{E}(X^2) = \int_0^1 x^2 \left(x + \frac{1}{2}\right) \mathrm{d}x = \frac{x^4}{4} + \frac{x^3}{6}\bigg|_0^1 = \frac{1}{4} + \frac{1}{6} = \frac{5}{12} = 0.417$$

As a check:

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

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

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

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

$$\mathrm{E}(XY) = \int_0^1 \int_0^1 xy(x + y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 \frac{x^2y^2}{2} + \frac{xy^3}{3}\bigg|_0^1 \mathrm{d}x = \int_0^1 \frac{x^2}{2} + \frac{x}{3}\,\mathrm{d}x$$
$$= \frac{x^3}{6} + \frac{x^2}{6}\bigg|_0^1 = \frac{1}{3} = 0.333$$

As a check:

```
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,

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

$$\rho = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} = \frac{-0.007}{\sqrt{0.076 \times 0.076}} = -0.0909$$

As a check:

```
-0.007/sqrt(.076^2)
```

```
## [1] -0.09210526
```

Using exact values:

```
(1/3-(7/12)^2)/sqrt((5/12-(7/12)^2)^2)
```

```
## [1] -0.09090909
```

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

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

$$\text{Var}(3X + 2Y) = \text{Var}(3X) + \text{Var}(2Y) + 2\text{Cov}(3X, 2Y) = 9\text{Var}(X) + 4\text{Var}(Y) + 12\text{Cov}(X, Y)$$

$$= 9 * 0.076 + 4 * 0.076 + 12 * -0.007 = 0.910$$

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 $\text{E}(X)$ and $\text{E}(Y)$.

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

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

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

$$\text{E}(X^2) = \int_0^1 x^2 \cdot 2x \, \text{d}x = \left.\frac{x^4}{2}\right|_0^1 = 0.5$$

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

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

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?

$$\text{E}(XY) = \int_0^1 \int_0^{2x} xy \, \text{d}y \, \text{d}x = \int_0^1 \left.\frac{xy^2}{2}\right|_0^{2x} \text{d}x = \int_0^1 2x^3 \, \text{d}x = \left.\frac{x^4}{2}\right|_0^1 = \frac{1}{2}$$

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 $\text{E}(XY)=\text{E}(X)\text{E}(Y)$.

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

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

Thus,

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

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	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{63}$
	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

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

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 roles.

```
(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

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

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

$$E[X|Y=2] = E[X] + 5$$

So

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

$$E[x] - \frac{2}{3}E[X] = \frac{2}{3} + \frac{3}{3} + \frac{5}{3}$$

$$\frac{1}{3}E[X] = \frac{10}{3}$$

$$E[X] = 10$$

6. ADVANCED: Let $X_1, X_2, ..., X_n$ 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, $\mathrm{E}(X_i) = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$).

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

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

$$\mathrm{E}(S) = \mathrm{E}(X_1 + X_2 + ... + X_n) = \mathrm{E}(X_1) + \mathrm{E}(X_2) + ... + \mathrm{E}(X_n) = \mu + \mu + ... + \mu = n\mu$$

Since the X_is are all independent:

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

$$\mathrm{E}(\bar{X}) = \frac{1}{n}\mathrm{E}(X_1 + X_2 + ... + X_n) = \frac{1}{n}n\mu = \mu$$
$$\mathrm{Var}(\bar{X}) = \frac{1}{n^2}\mathrm{Var}(X_1 + X_2 + ... + X_n) = \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$$

File Creation Information

- File creation date: 2020-10-21
- 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