Multivariate Distribution Applications Solutions

Lt Col Ken Horton   Lt Col Kris Pruitt   Professor Bradley Warner

29 September, 2020

Exercises

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

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

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

   a. Verify that f is a valid pdf.

$$\int_0^1 \int_0^1 x + y \, \mathrm{d}y \, \mathrm{d}x = \int_0^1 xy + \frac{y^2}{2} \bigg|_0^1 \mathrm{d}x = \int_0^1 x + \frac{1}{2} \mathrm{d}x = \frac{x^2}{2} + \frac{x}{2} \bigg|_0^1 = 1$$

Or

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

```
## $integral
## [1] 1
##
## $error
## [1] 0
##
## $functionEvaluations
## [1] 17
##
## $returnCode
## [1] 0
```

   b. Find the marginal pdfs of X and Y.

$$f_X(x) = \int_0^1 x + y \, \mathrm{d}y = xy + \frac{y^2}{2} \bigg|_0^1 = x + \frac{1}{2}$$

where $0 \leq x \leq 1$.

Similarly, $f_Y(y) = y + \frac{1}{2}$ for $0 \leq y \leq 1$.

c. Find the conditional pdfs of $X|Y = y$ and $Y|X = x$.

$$f_{X|Y=y}(x) = \frac{x+y}{y+\frac{1}{2}}$$

where $0 \le x \le 1$.

Similarly, $f_{Y|X=x}(y) = \frac{x+y}{x+\frac{1}{2}}$ for $0 \le y \le 1$.

d. Find the following probabilities: $P(X < 0.5)$; $P(Y > 0.8)$; $P(X < 0.2, Y \ge 0.75)$; $P(X < 0.2|Y \ge 0.75)$; $P(X < 0.2|Y = 0.25)$; Optional - $P(X \le Y)$.

$$P(X < 0.5) = \int_0^{0.5} x + \frac{1}{2}\, dx = \frac{x^2}{2} + \frac{x}{2}\Big|_0^{0.5} = 0.375$$

```
integrate(function(x)(x+1/2),0,1/2)
```

```
## 0.375 with absolute error < 4.2e-15
```

Or using multivariate integration, integrate out y.

```
adaptIntegrate(f, lowerLimit = c(0, 0), upperLimit = c(1/2, 1))
```

```
## $integral
## [1] 0.375
##
## $error
## [1] 5.551115e-17
##
## $functionEvaluations
## [1] 17
##
## $returnCode
## [1] 0
```

$$P(Y < 0.8) = \int_{0.8}^1 y + \frac{1}{2}\, dy = \frac{y^2}{2} + \frac{y}{2}\Big|_{0.8}^1 = 1 - 0.72 = 0.28$$

```
adaptIntegrate(f, lowerLimit = c(0, 0.8), upperLimit = c(1, 1))
```

```
## $integral
## [1] 0.28
##
## $error
## [1] 5.551115e-17
##
## $functionEvaluations
## [1] 17
##
## $returnCode
## [1] 0
```

$$P(X < 0.2, Y \geq 0.75) = \int_0^{0.2} \int_{0.75}^1 x + y \, dy \, dx = \int_0^{0.2} xy + \frac{y^2}{2} \bigg|_{0.75}^1 dx$$

$$= \int_0^{0.2} x + \frac{1}{2} - \frac{3x}{4} - \frac{9}{32} \, dx = \int_0^{0.2} \frac{x}{4} + \frac{7}{32} \, dx = \frac{x^2}{8} + \frac{7x}{32} \bigg|_0^{0.2} = 0.04875$$

```
adaptIntegrate(f, lowerLimit = c(0, 0.75), upperLimit = c(0.2, 1))
```

```
## $integral
## [1] 0.04875
##
## $error
## [1] 0
##
## $functionEvaluations
## [1] 17
##
## $returnCode
## [1] 0
```

$$P(X < 0.2 | Y \geq 0.75) = \frac{P(X < 0.2, Y \geq 0.75)}{P(Y \geq 0.75)} = \frac{0.06875}{\int_{0.75}^1 y + \frac{1}{2} \, dy} = \frac{0.06875}{0.34375} = 0.2$$

For

$$P(X < 0.2 | Y = 0.25)$$

we need

$$f_{X|Y=.25}(x) = \frac{x+y}{y+\frac{1}{2}} \bigg|_{y=0.25} = \frac{x+.25}{.25+\frac{1}{2}} = \frac{x+.25}{.75} = \frac{4x+1}{3}$$

$$P(X < 0.2 | Y = 0.25) = \int_0^{0.2} \frac{4x+1}{3} \, dx$$

$$= \frac{1}{3} \left(2x^2 + x\right) \bigg|_0^{0.2} = \frac{1}{3} \left(2 \cdot 0.2^2 + 0.2\right) = 0.0084$$

```
f2 <- function(x) { (2*x[1]^2 + x[1])/3 } # "x" is vector
adaptIntegrate(f2, lowerLimit = c(0), upperLimit = c(0.2))
```

```
## $integral
## [1] 0.008444444
##
## $error
## [1] 9.375217e-17
##
## $functionEvaluations
## [1] 15
##
## $returnCode
## [1] 0
```

Optional

$$P(X \leq Y) = \int_0^1 \int_0^x x + y \, \mathrm{d}y \, \mathrm{d}x = \int_0^1 xy + \left.\frac{y^2}{2}\right|_0^x \mathrm{d}x = \int_0^1 \frac{3x^2}{2} \, \mathrm{d}x = \left.\frac{x^3}{2}\right|_0^1 = \frac{1}{2}$$

2. In the Notes, we saw an example where $f_X(x) = f_{X|Y=y}(x)$ and $f_Y(y) = f_{Y|X=x}(y)$. This is not common and is important. What does this imply about X and Y?

Since the conditional density function is always equal to the marginal, it means that X and Y are independent of one another. Also, if the conditioning variable does not appear in the conditional density function and the domain of the joint density is rectangular, the bounds of the two variables are constants, the random variables are independent. The variables in the previous problem are dependent, look at the conditional density functions to see that the conditional density depends on the conditioned variable.

3. ADVANCED: Recall on an earlier assignment, we came up with random variables to describe timeliness at an airport. Suppose over the course of 210 days, on each day we recorded the number of customer complaints regarding timeliness. Also on each day, we recorded the weather (our airport is located somewhere without snow and without substantial wind). The data are displayed below.

		Weather Status	
	Clear	Light Rain	Rain
0	28	11	4
1	18	15	8
2	17	25	12
num complaints 3	13	15	16
4	8	8	10
5	0	1	1

First, define two random variables for this scenario. One of them (# of complaints) is essentially already a random variable. For the other (weather status) you will need to assign a number to each status.

a. Use the table above to build an empirical joint pmf of the two random variables.

We will simply label the weather random variable as 0, 1, 2. We convert to probabilities by dividing by 210.

		X : Weather Status	
	0 (Clear)	1 (Light Rain)	2 (Rain)
0	0.133	0.052	0.019
1	0.086	0.071	0.038
2	0.081	0.119	0.057
Y: num complaints 3	0.062	0.071	0.076
4	0.038	0.038	0.048
5	0	0.005	0.005

b. Find the marginal pmfs of each random variable.

$$f_X(x) = \begin{cases} 0.400, & x = 0 \\ 0.357, & x = 1 \\ 0.243, & x = 2 \\ 0, & \text{otherwise} \end{cases}$$

4

$$f_Y(y) = \begin{cases} 0.205, & y = 0 \\ 0.195, & y = 1 \\ 0.257, & y = 2 \\ 0.210, & y = 3 \\ 0.124, & y = 4 \\ 0.010, & y = 5 \\ 0, & \text{otherwise} \end{cases}$$

c. Find the probability of fewer than 3 complaints.

$$P(Y < 3) = 0.205 + 0.195 + 0.257 = 0.657$$

d) Find the probability of fewer than 3 complaints given there is no rain.

$$P(Y < 3|X = 0) = \frac{0.133 + 0.086 + 0.081}{0.657} = 0.457$$

Optional for those of you that like Calc III and want a challenge.

4. Let X and Y be continuous random variables with joint pmf:

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

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

a. Verify that f is a valid pdf.

$$\int_0^1 \int_0^{2x} 1 \, dy \, dx = \int_0^1 y \Big|_0^{2x} \, dx = \int_0^1 2x \, dx = x^2 \Big|_0^1 = 1$$

b. Find the marginal pdfs of X and Y.

$$f_X(x) = \int_0^{2x} 1 \, dy = y \Big|_0^{2x} = 2x$$

where $0 \le x \le 1$.

$$f_Y(y) = \int_{y/2}^1 1 \, dx = x \Big|_{y/2}^1 = 1 - \frac{y}{2}$$

where $0 \le y \le 2$.

c. Find the conditional pdfs of $X|Y = y$ and $Y|X = x$.

$$f_{X|Y=y}(x) = \frac{1}{1 - \frac{y}{2}} = \frac{2}{2 - y}$$

where $y/2 \le x \le 1$.

$$f_{Y|X=x}(y) = \frac{1}{2x}$$

where $0 \le y \le 2x$.

d. Find the following probabilities: $P(X < 0.5)$; $P(Y > 1)$; $P(X < 0.5, Y \le 0.8)$; $P(X < 0.5|Y = 0.8)$; Optional $P(Y \le 1 - X)$. (It would probably help to draw some pictures.)

$$P(X < 0.5) = \int_0^{0.5} 2x \, dx = x^2 \Big|_0^{0.5} = 0.25$$

$$P(Y > 1) = \int_1^2 1 - \frac{y}{2} \, dy = y - \frac{y^2}{4} \Big|_1^2 = 1 - \frac{3}{4} = 0.25$$

$$P(X < 0.5, Y \le 0.8) = \int_0^{0.4} \int_0^{2x} 1 \, dy \, dx + \int_{0.4}^{0.5} \int_0^{0.8} 1 \, dy \, dx = 0.16 + 0.08 = 0.24$$

$$P(X < 0.5|Y = 0.8) = \int_{0.4}^{0.5} \frac{2}{2 - 0.8} \, dx = \frac{5x}{3} \Big|_{0.4}^{0.5} = 0.1667$$

$$P(Y \le 1 - X) = \int_0^{1/3} \int_0^{2x} 1 \, dy \, dx + \int_{1/3}^1 \int_0^{1-x} 1 \, dy \, dx = \int_0^{1/3} 2x \, dx + \int_{1/3}^1 1 - x \, dx$$

$$= \frac{1}{9} + \left(x - \frac{x^2}{2} \right)_{1/3}^1 = \frac{1}{9} + \frac{1}{2} - \frac{1}{3} + \frac{1}{18} = \frac{1}{3}$$

File Creation Information

- File creation date: 2020-09-29
- 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
- `cubature` package version: 2.0.4.1
- `mosaicCalc` package version: 0.5.1