Unit 5 Homework: Joint Distributions

Tako Hisada

10/7/2017

1. Unladen Swallows

```
library(MASS)
set.seed(898)
```

We know SD of Wingspan is 4. Since $V(X) = \sqrt(\sigma_X^2)$, $V(W) = 4^2 = 16$.

$$V(V) = V[h(W)] = V[\frac{1}{2}w + U] = a^2 * \sigma_W^2 + V(U) = (\frac{1}{2})^2 * 4^2 + 1 = 5$$

Now we want to compute $Cov(W, V)$.

$$Cov(W, V) = Cov(W, \frac{1}{2}w + U) = Cov(W, \frac{1}{2}W) + Cov(W, U) = Cov(W, 1/2W) + 0$$

$$Cov(W, 1/2W) = E((W - E(W)(\frac{1}{2}W - E(\frac{1}{2}W)) = E((W - E(W)(\frac{1}{2}W - \frac{1}{2}E(W))$$

$$E((W - E(W)\frac{1}{2}(W - E(W)) = \frac{1}{2}E((W - E(W))^2) = \frac{1}{2}VAR(W) = \frac{1}{2} * 16 = 8$$

```
# Wingspan
w = rnorm(100, mean = 10, sd = 4)
mu_w = 10
sd_w = 4
v_w = sd_w^2
a_w = 1/2

# Velocity
v = a_w * w + rnorm(100)
v_v = a_w^2 * sd_w^2 + 1

# Covariance
cov_w_v = 1/2*v_w

# Covariance-Variance Matrix
(Sigma = matrix(c(v_w, cov_w_v, cov_w_v, v_v), 2, 2))
```

```
##      [,1] [,2]
## [1,]   16    8
## [2,]    8    5
```

2. Broken Rulers

a. Conditional expectation of Y given X, E(Y|X)

X is uniformly distributed.

$$f(Y|X) = \frac{1}{x}$$

$$E(Y|X) = \int_0^X yf(Y|X)dy$$

$$E(Y|X) = \int_0^X y * \frac{1}{x} = \frac{y^2}{x}|_0^X = \frac{x}{2}$$

b. Unconditional expectation of Y

$$E[Y] = E[E[Y|X]]$$

We know from a E[Y|X] = $\frac{x}{2}$

$$E[Y] = E[\frac{x}{2}] = \frac{1}{2}E(X)$$

$$E[X] = \int_0^1 xf(x)dx = \frac{x^2}{2}|_0^1 = \frac{1}{2}$$

$$E[Y] = \frac{1}{2}E[X] = \frac{1}{2} * \frac{1}{2} = \frac{1}{4}$$

c. Compute E(XY)

$$E[XY] = E[E(XY|X)] = E[X * E(Y|X)] = E[X\frac{X}{2}] = E[\frac{X^2}{2}] = \frac{1}{2}E[X^2] = \frac{1}{2}E[X^2]$$

$$E[X^2] = E[h(X)] = \int_0^1 x^2 f(x)dx = \int_0^1 x^2 dx = \frac{x^3}{3}|_0^1 = \frac{1}{3}$$

$$E[XY] = \frac{1}{2}E[X^2] = \frac{1}{2} * \frac{1}{3} = \frac{1}{6}$$

d. Compute Cov(X, Y)

$$Cov(X, Y) = E(XY) - \mu_X \mu_Y = \frac{1}{6} - \frac{1}{2} * \frac{1}{4} = \frac{1}{24}$$

```
(q2d = 1/6-1/2*1/4)
```

```
## [1] 0.04166667
```

3. Great Time to Watch Async

a. Total expected waiting time when you take the Caltrain each morning and each evening for 5 days in a row

```
i <- 0:5
ex = sum(1/6*i)
j <- 0:10
ey = sum(1/11*j)
exy = ex + ey
```

$$E(X+Y) = E(X) + E(Y)$$

$$E(X) = \sum_{x=0}^{5} x * p(x) = 0 * \frac{1}{6} + 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} = 2.5$$

$$E(Y) = \sum_{x=0}^{10} y * p(y) = 0*\frac{1}{10}+1*\frac{1}{10}+2*\frac{1}{10}+3*\frac{1}{10}+4*\frac{1}{10}+5*\frac{1}{10}+6*\frac{1}{10}+7*\frac{1}{10}+8*\frac{1}{10}+9*\frac{1}{10}+10*\frac{1}{10} = 5$$

$$E(X+Y) = E(X) + E(Y) = 2.5 + 5 = 7.5$$

Since we are being asked for the total expected waiting time when we took the Caltrain for 5 days in a row, we need to multiply this value by 5.

```
q3a = exy*5
```

$$E(X+Y) = 7.5 * 5 = 37.5$$

37.5 minutes

b. Variance of the total waiting time

$$V(X+Y) = V(X) + V(Y) + 2 * COV(X+Y)$$

$$V(X) = E(X^2) - E(X)^2$$

$$V(Y) = E(Y^2) - E(Y)^2$$

```
compExUniform <- function(x) {
  sum = 0
  for(i in 0:x) {
    sum = sum + i * 1/(x+1)
  }
  sum
}

ex = compExUniform(5)
ey = compExUniform(10)

compExUniformSq <- function(x) {
  sum = 0
  for(i in 0:x) {
```

```
    sum = sum + i^2 * 1/(x+1)
  }
  sum
}
exx_sq <- compExUniformSq(5)
exy_sq <- compExUniformSq(10)

varx <- exx_sq - ex^2
vary <- exy_sq - ey^2
```

$$Cov(X,Y) = E(XY) - \mu_X \mu_Y$$

```
compExXYUniform <- function(x, mux, y, muy) {
  sum = 0
  for(i in 0:x) {
    for(j in 0:y) {
      sum = sum + ((i-mux)*(j-muy))/((x+1)*(y+1))
    }
    #print(sprintf("END SUM: %f", sum))
  }
  sum
}
exxy <- compExXYUniform(5, 2.5, 10, 5)

cov_xy = exxy - ex*ey
```

Now we can compute $V(XY)$

```
varxy = varx + vary - 2*cov_xy
```

$$37\frac{11}{12}$$

However we need to multiple this by 5 as we are interested in finding out the variance of the total waiting time

```
q2b = 5^2 * varxy
```

$$947\frac{11}{12}$$

c. Expected value of the difference between the total evening waiting time and the total morning waiting time over all 5 days

$$E(X - Y)$$

```
i <- 0:5
ex = sum(1/6*i)
j <- 0:10
ey = sum(1/11*j)
exy_diff = abs(ex - ey)
```

$$E(X-Y) = E(X) - E(Y)$$

$$E(X) = \sum_{x=0}^{5} x * p(x) = 0 * \frac{1}{6} + 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} = 2.5$$

$$E(Y) = \sum_{x=0}^{10} y * p(y) = 0 * \frac{1}{10} + 1 * \frac{1}{10} + 2 * \frac{1}{10} + 3 * \frac{1}{10} + 4 * \frac{1}{10} + 5 * \frac{1}{10} + 6 * \frac{1}{10} + 7 * \frac{1}{10} + 8 * \frac{1}{10} + 9 * \frac{1}{10} + 10 * \frac{1}{10} = 5$$

$$E(X-Y) = |E(X) - E(Y)| = 2.5$$

Since we are being asked for the total expected waiting time when we took the Caltrain for 5 days in a row, we need to multiply this value by 5.

```
q3c = exy_diff*5
```

$$E(X+Y) = 2.5 * 5 = 12.5$$

12.5 minutes

d. Variance of the difference between the total evening time and the total morning waiting time over all 5 days

$$V(X-Y) = V(X) + V(Y)$$

$$V(X) = E(X^2) - E(X)^2$$

$$V(Y) = E(Y^2) - E(Y)^2$$

```
compExUniform <- function(x) {
  sum = 0
  for(i in 0:x) {
    sum = sum + i * 1/(x+1)
  }
  sum
}

ex = compExUniform(5)
ey = compExUniform(10)

compExUniformSq <- function(x) {
  sum = 0
  for(i in 0:x) {
    sum = sum + i^2 * 1/(x+1)
  }
  sum
}
exx_sq <- compExUniformSq(5)
exy_sq <- compExUniformSq(10)

varx <- exx_sq - ex^2
vary <- exy_sq - ey^2
```

Now we can compute $V(XY)$

```
varxy = varx + vary
```

$$12\frac{11}{12}$$

However we need to multiple this by 5 as we are interested in finding out the variance of the total waiting time

```
q3d = 5^2 * varxy
```

$$322\frac{11}{12}$$

4. Maximizing Correlation

Show that if $Y = aX + b$ where X and Y are random variables and a != 0, corr(X, Y) = -1 or +1

$$Corr(X, Y) = P_{X,Y} = \frac{Cov(X, Y)}{\sigma_X * \sigma_Y}$$
$$Corr(X, aX + b) = P_{X,aX+b} = \frac{Cov(X, aX + b)}{\sigma_X * \sigma_{aX+b}}$$

Since adding a constant to random variables does not change their covariances or variances, we can say:

$$\frac{Cov(X, aX)}{\sigma_X * \sigma_{aX}}$$

We can also move the constant a out of Cov and σ

$$\frac{a * Cov(X, X)}{|a| * \sigma_X * \sigma_X}$$

We can cancel the constant a

$$\frac{Cov(X, X)}{\sigma_X * \sigma_X} = \frac{Var(X)}{Var(X)} = 1$$

In case a is a negative number, however, -1 remains in the numerator which makes Corr(X, Y) = -1.

$$\frac{-Cov(X, X)}{\sigma_X * \sigma_X} = \frac{Var(X)}{Var(X)} = -1$$

“‘