Calculating covariance of joint probability mass function in R

Asked 5 years ago Modified 5 years ago Viewed 2k times Part of R Language Collective

I have a joint probability mass function of two variables X,Y like here

		r		
	0	1	2	$f_G(g)$
0	$\frac{\binom{4}{3}}{\binom{9}{3}} = \frac{4}{84}$	$\frac{\binom{2}{1}\binom{4}{2}}{\binom{9}{3}} = \frac{12}{84}$	$\frac{\binom{2}{2}\binom{4}{1}}{\binom{9}{3}} = \frac{4}{84}$	$\frac{20}{84}$
1	$\frac{\binom{3}{1}\binom{4}{2}}{\binom{9}{3}} = \frac{18}{84}$	$\frac{\binom{2}{1}\binom{3}{1}\binom{4}{1}}{\binom{9}{3}} = \frac{24}{84}$	$\frac{\binom{2}{2}\binom{3}{1}}{\binom{9}{3}} = \frac{3}{84}$	$\frac{45}{84}$
g 2	$\frac{\binom{3}{2}\binom{4}{1}}{\binom{9}{3}} = \frac{12}{84}$	$\frac{\binom{2}{1}\binom{3}{2}}{\binom{9}{3}} = \frac{6}{84}$	0	$\frac{18}{84}$
3	$\frac{\binom{3}{3}}{\binom{9}{3}} = \frac{1}{84}$	0	0	$\frac{1}{84}$
$f_R(r)$	$\frac{35}{84}$	$\frac{42}{84}$	$\frac{7}{84}$	

How can I calculate the covariance in R?

I created two vectors x,y and fed them into cov(), but I get the wrong result.

How can I do this right?

Thanks in advance and happy coding!

r statistics probability Edit tags

Share Edit Follow Close Flag

edited May 31, 2018 at 13:44
Maurits Evers
49.3k 4 46 68

asked May 31, 2018 at 10:51
nequalstim
3 1

This is probably better at home on CrossValidated, the statistics website on the SE network. – JAD May 31, 2018 at 11:13

ah thanks, I'll ask there then! – nequalstim May 31, 2018 at 11:14

As far as an R implementation is concerned, please take a look my step-by-step answer below. – Maurits Evers May 31, 2018 at 12:56

1 Answer

Sorted by: Reset to default Date modified (newest first)

Since SO is a coding forum, I'll leave working out the math/stats details up to you. Here is an implementation in R.

1. We start by noting the sample spaces for X and Y

```
# For G
G <- 0:3;

# For R
R <- 0:2;
```

2. The joint probability mass function is given by the following matrix

```
joint_pmf <- matrix(
    c(4/84, 12/84, 4/84,
      18/84, 24/84, 3/84,
      12/84, 6/84, 0,
```

```
        1/84, 0, 0),
      ncol = 3, byrow = T);
```

3. We calculate the population means

```
# For G
mu_G <- rowSums(joint_pmf) %*% G;

# For R
mu_R <- colSums(joint_pmf) %*% R;
```

4. We can make use of the theorem `Cov(X, Y) = E[XY] - E[X]E[Y]` to calculate the covariance

```
cov_GR <- G %*% joint_pmf %*% R - mu_G * mu_R;
#           [,1]
#[1,] -0.1666667
```

where we have used the fact that `E[G] = mu_G` and `E[R] = mu_R` are the respective population means.

Share Edit Follow Flag

edited May 31, 2018 at 12:51

answered May 31, 2018 at 12:32

Maurits Evers

49.3k 4 46 68