Mathematics Stack Exchange is a question and answer site for people studying math at any level and professionals in related fields. It only takes a minute to sign up.

Sign up to join this community

Anybody can ask a question

Anybody can answer

The best answers are voted up and rise to the top



Find the covariances of a multinomial distribution

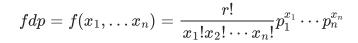
Asked 4 years, 1 month ago Active 1 year, 11 months ago Viewed 8k times



If (X_1, \dots, X_n) is a vector with multinomial distribution, proof that $Cov(X_i, X_j) = -rp_i p_j$, $i \neq j$ where r is the number of trials of the experiment, p_i is the probability of success for the variable X_i .



8





$$\text{if } x_1+x_2+\cdots+x_n=r$$

I'm trying to use the property: $Cov(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j]$ and find that $E[X_i] = rp_i$, but I don't know the efficient way to calculate $E[X_i X_j]$.

probability combinatorics statistics

edited Feb 24 '16 at 1:46

asked Feb 24 '16 at 1:18



What is $E(X_i X_i)$ for r = 1? – A.S. Feb 24 '16 at 1:37 \nearrow

By using our site, you acknowledge that you have read and understand our Cookie Policy, Privacy Policy, and our Terms of Service.



X

As what A.S. hinted, one common trick is to express $X_i = \sum_{k=1}^r Y_{i,k}$, $X_j = \sum_{l=1}^r Y_{j,l}$ and use linearity of covariance. By independence across different multinomial trials, you only left the calculate the case with $Cov[Y_{i,k}, Y_{j,k}]$. But those Y are indicators only (i.e. the r=1 case mentioned by A.S.) which is easy to calculate. – BGM Feb 24 '16 at 3:43

1 Answer





13

We can use indicator random variables to help simplify the covariance expression. We can interpret the problem as r independent rolls of an n sided die. Let X_i be the number of rolls that result in side i facing up, and let $I_k^{(i)}$ be an indicator equal to 1 when roll k is equal to i and 0 otherwise. Then, we can express X_i and X_i as follows:



$$X_i = \sum_{k=1}^r I_k^{(i)} \ \ ext{and} \ \ X_j = \sum_{k=1}^r I_k^{(j)}$$

Let's re-write the covariance using indicators:

$$Cov(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j]$$

Let's compute the first term:

$$egin{aligned} E[X_i X_j] &= E\Big[(\sum_{k=1}^r I_k^{(i)})(\sum_{l=1}^r I_l^{(j)})\Big] = \sum_{k=l} Eig[I_k^{(i)} I_l^{(j)}ig] + \sum_{k
eq l} Eig[I_k^{(i)} I_l^{(j)}ig] = \ &= 0 + \sum_{k
eq l} Eig[I_k^{(i)}ig] Eig[I_l^{(j)}ig] = \sum_{k
eq l} p_i p_j = (r^2 - r) p_i p_j \end{aligned}$$

where we expanded the product of sums, used linearity of expectation and the fact that when k=l we can't simultaneously roll i and j on the same trial k=l (making the product of indicators zero), finally we applied independence of rolls that enabled us to write it as a product of probabilities. Let's compute the remaining term:

$$E[X_i] = E[\sum_{k=1}^r I_k^{(i)}] = \sum_{k=1}^r E[I_k^{(i)}] = rp_i$$

Therefore, the covariance equals:

$$\mathrm{Cov}(X_i, X_j) = E[X_i X_j] - E[X_i] E[X_j] = (r^2 - r) p_i p_j - r^2 p_i p_j = -r p_i p_j$$

Notice that $Cov(X_i, X_j) = -rp_i p_j < 0$ is negative, this makes sense intuitively since for a fixed number of rolls r, if we roll many outcomes i, this reduces the number of possible outcomes j, and therefore X_i and X_j are negatively correlated!

answered Apr 3 '18 at 21:41



By using our site, you acknowledge that you have read and understand our Cookie Policy, Privacy Policy, and our Terms of Service.

