$\boxed{\text{MULTIVARIATE RANDOM VARIABLES}}$

Mauro Diaz

$\boxed{8.2}$

Exercise 8.2 (Multinomial expected value) Find the expected value, variance and covariance of the multinomial distribution. Hint: First find the expected value for $n = 1$ and then use the fact that the trials are independent.

$$p(x) = \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k} \quad \text{where} \quad n = \sum_{i=1}^{k} x_i$$

$n \equiv$ number of trials

$X_i \equiv$ event with prob. $p_i^{x_i}$.

We will fix $n=1$. $X'_1 \equiv X_1$ with $n=1$

$$\mathbb{E}[X'_1] = \sum_{x_1=0}^{1} \cdots \sum_{x_k=0}^{1} \frac{1}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k} x_1 = \sum_{x_1=0}^{1} \frac{p_1^{x_1} x_1}{x_1!} \underbrace{\sum_{x_2=0}^{1} \cdots \sum_{x_k=0}^{1} \frac{1}{x_2! \cdots x_k!} p_2^{x_2} \cdots p_k^{x_k}}_{\text{"}1\text{"}} =$$

$$= 0 \cdot 1 + p_1 \cdot 1 = p_1$$

We can do the analogous computation for every component X_i of $X = (X_1, \dots, X_k)$ and we have $\mathbb{E}[X'_i] = p_i$ when n is fixed to 1.

Now, we can write $X = \sum_{j=1}^{n} Y_j$ where $X = (X_1, \dots, X_k)$ with $n \in \mathbb{N}$ and Y_j are multinomial distr. with $n=1$ fixed.

$$\mathbb{E}[X] = \mathbb{E}\left[\sum Y_j\right] = \sum \mathbb{E}[Y_j] = \sum_{j=1}^{n} (p_1, \dots, p_k) = (np_1, \dots, np_k)$$

For the variance, we need to compute $\mathbb{E}[X^2]$. We can do a same computation that for the mean: fix $n=1$.

$$\mathbb{E}[X'^2_1] = \sum_{x_1=0}^{1} \cdots \sum_{x_k=0}^{1} \frac{1}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k} x_1^2 = \sum_{x_1=0}^{1} \frac{p_1^{x_1} x_1^2}{x_1!} \underbrace{\sum_{x_2=0}^{1} \cdots \sum_{x_k=0}^{1} \frac{1}{x_2! \cdots x_k!} p_2^{x_2} \cdots p_k^{x_k}}_{\text{"}1\text{"}} =$$

$$= 0 \cdot 1 + p_1 \cdot 1 = p_1 \qquad \Rightarrow \qquad \mathbb{E}[X^2] = \sum_{j=1}^{n} \mathbb{E}[Y_j^2] = (np_1, \dots, np_k)$$

Combining all: $\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = (np_1(1-p_1), \dots, np_k(1-p_k))$

Finally, for the covariance, we need to calculate $\mathbb{E}[X_i X_j]$.

We will apply the same idea. We start fixing $n=1$. Without loss of generality, we will assume $i=1$ and $j=2$, and could be later generalized.

$$\mathbb{E}[X_1'X_2'] = \sum_{x_1=0}^{1}\sum_{x_2=0}^{1} \frac{p_1^{x_1}\,x_1}{x_1!}\,\frac{p_2^{x_2}\,x_2}{x_2!}\,\underbrace{\sum_{x_3=0}^{1}\cdots\sum_{x_k=0}^{1}\frac{1}{x_3!\cdots x_k!}\,p_3^{x_3}\cdots p_k^{x_k}}_{1} =$$

$$= \sum_{x_1=0}^{1}\sum_{x_2=0}^{1} \frac{p_1^{x_1}\,x_1}{x_1!}\,\frac{p_2^{x_2}\,x_2}{x_2!} = 0$$

$\sum x_i = n \iff x_1+x_2 = 1 \iff x_1=0 \lor x_2=0$

So, now, in the covariance formula:

$$\mathrm{Cov}(X_i', X_j') = \mathbb{E}[X_i'X_j'] - \mathbb{E}[X_i']\,\mathbb{E}[X_j'] = -p_ip_j$$

Now, for $n\in\mathbb{N}$, let as always $X = \sum_{j=1}^{n} Y_j \Rightarrow X_1 = \sum_{j=1}^{n} X_{1,j}'$, $X_2 = \sum_{j=1}^{n} X_{2,j}'$

$$\Rightarrow \mathrm{Cov}(X_1, X_2) = \mathbb{E}[X_1X_2] - \mathbb{E}[X_1]\,\mathbb{E}[X_2] =$$

$$= \mathbb{E}\left[\sum_{j=1}^{n} X_{1,j}' \sum_{j=1}^{n} X_{2,j}'\right] - n^2p_1p_2 =$$

$j=l \Rightarrow 0$
$j\neq l \Rightarrow p_1p_2$

$$= \sum_{j=1}^{n}\sum_{l=1}^{n} \mathbb{E}[X_{1,j}'X_{2,j}'] - n^2p_1p_2$$

$$= (n^2-n)\,p_1p_2 - n^2p_1p_2 = -np_1p_2$$