1. A Multivariate Discrete Distribution – The Multinomial Distribution

Remark 1.1. While considering a Bernoulli or a Binomial RV, we looked at random experiments with exactly two outcomes. We now consider random experiments with two or more than two outcomes. Suppose a random experiment terminates in one of k possible outcomes $A_1, A_2, \cdots, A_k$ for $k \geq 2$. More generally, we may also consider random experiments which terminate in one of k mutually exclusive and exhaustive events $A_1, A_2, \cdots, A_k$ with $k \geq 2$. Write $p_j = \mathbb{P}(A_j), j = 1, 2, \cdots, k$, which does not change from trial to trial. Then, $p_1 + p_2 + \cdots + p_k = 1$. Suppose n trials are conducted independently and let $X_j, j = 1, 2, \cdots, k$ denote the number of times event A_j has occured, respectively. Then the RVs $X_1, X_2, \cdots, X_k$ satisfy the relation $X_1 + X_2 + \cdots + X_k = n$ and we have

$$\mathbb{P}(X_1 = x_1, \cdots, X_k = x_k) = \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k}$$

for $x_1, \cdots, x_k \in \{0, 1, \cdots, n\}$ with $x_1 + \cdots + x_k = n$. The probability is zero otherwise. Removing the redundancy we have the joint p.m.f. of $(X_1, X_2, \cdots, X_{k-1})$ given by

$$\begin{aligned} & f_{X_1, \cdots, X_{k-1}}(x_1, \cdots, x_{k-1}) \\ &= \frac{n!}{x_1! \cdots x_{k-1}!(n - x_1 - \cdots - x_{k-1})!} p_1^{x_1} \cdots p_{k-1}^{x_{k-1}} (1 - p_1 - \cdots - p_{k-1})^{n - x_1 - \cdots - x_{k-1}} \end{aligned}$$

for $x_1, \cdots, x_k \in \{0, 1, \cdots, n\}$ with $x_1 + \cdots + x_{k-1} \leq n$ and zero otherwise.

Example 1.2 (Multinomial Distribution)**.** A random vector $X = (X_1, \cdots, X_{k-1})$ is said to follow the Multinomial distribution with parameters n and $p_1, p_2, \cdots, p_k$ if the joint p.m.f. is as in Remark 1.1 above. We now list some properties of the multinomial distribution.

(a) We first compute the joint MGF. For $t_1, t_2, \cdots, t_{k-1} \in \mathbb{R}$,

$$\begin{aligned} & M_X(t_1, t_2, \cdots, t_{k-1}) \\ &= \mathbb{E} \exp(t_1 X_1 + t_2 X_2 + \cdots + t_{k-1} X_{k-1}) \\ &= \sum_{\substack{x_1, \cdots, x_k \in \{0, 1, \cdots, n\} \\ x_1 + \cdots + x_{k-1} \leq n}} \frac{n! \exp(t_1 x_1 + t_2 x_2 + \cdots + t_{k-1} x_{k-1})}{x_1! \cdots x_{k-1}!(n - x_1 - \cdots - x_{k-1})!} p_1^{x_1} \cdots p_{k-1}^{x_{k-1}} p_k^{n - x_1 - \cdots - x_{k-1}} \end{aligned}$$

$$= \sum_{\substack{x_1,\cdots,x_k\in\{0,1,\cdots,n\}\\ x_1+\cdots+x_{k-1}\le n}} \frac{n!}{x_1!\cdots x_{k-1}!(n-x_1-\cdots-x_{k-1})!}\left(p_1e^{t_1}\right)^{x_1}\cdots\left(p_{k-1}e^{t_{k-1}}\right)^{x_{k-1}} p_k^{n-x_1-\cdots-x_{k-1}}$$

$$= \left(p_1e^{t_1}+p_2e^{t_2}+\cdots+p_{k-1}e^{t_{k-1}}+p_k\right)^n$$

(b) If $t=(t_1,0,\cdots,0)\in\mathbb{R}^{k-1}$, then $M_X(t)=\mathbb{E}\exp(t_1X_1)=M_{X_1}(t_1)$. But, using the above expression for the joint MGF, we have $M_{X_1}(t_1)=M_X(t)=(p_1e^{t_1}+p_2+\cdots+p_{k-1}+p_k)^n=(p_1e^{t_1}+1-p_1)^n$. Therefore, $X_1\sim Binomial(n,p_1)$. Similarly, $X_i\sim Binomial(n,p_i),\forall i=2,\cdots,k-1$. In particular, $\mathbb{E}X_i=np_i, Var(X_i)=np_i(1-p_i)$.

(c) For distinct indices $i,j\in\{1,2,\cdots,k-1\}$,

$$M_{X_i,X_j}(t_i,t_j)=M_X(0,\cdots,0,t_i,0,\cdots,0,t_j,0,\cdots,0)=\left(p_ie^{t_i}+p_je^{t_j}+1-p_i-p_j\right)^n, \forall(t_i,t_j)\in\mathbb{R}^2.$$

Therefore (X_i,X_j) follows the trinomial distribution with the parameters $p_i,p_j,1-p_i-p_j$, i.e. multinomial distribution with the parameters $n=3$ and $p_i,p_j,1-p_i-p_j$.

(d) For distinct indices $i,j\in\{1,2,\cdots,k-1\}$, consider $t_i=t_j=t\in\mathbb{R}$. Then,

$$M_{X_i+X_j}(t)=M_{X_i,X_j}(t,t)=\left[(p_i+p_j)\,e^t+1-(p_i+p_j)\right]^n,$$

which shows $X_i+X_j\sim Binomial(n,p_i+p_j)$. Then $Var(X_i+X_j)=n(p_i+p_j)(1-p_i-p_j)$. Using the relation

$$Var(X_i+X_j)=Var(X_i)+Var(X_j)+2Cov(X_i,X_j),$$

we have $Cov(X_i,X_j)=-np_ip_j$. Consequently, the correlation between X_i and X_j is

$$\rho(X_i,X_j)=\frac{Cov(X_i,X_j)}{\sqrt{Var(X_i)\,Var(X_j)}}=-\left(\frac{p_ip_j}{(1-p_i)(1-p_j)}\right)^{\frac{1}{2}}.$$