Module 31

Some Special Multivariate Distributions

Multinomial Coefficients

- Let $k, n_1, \ldots, n_{k-1}$ and n be non-negative integers such that $k \geq 2$, $\sum_{i=1}^{k-1} n_i \leq n$. Consider a collection of n items comprising of

$$n_1 \text{ identical items of type } 1$$

$$n_2 \text{ identical items of type } 2$$

$$\vdots$$

$$n_{k-1} \text{ identical items of type } k-1$$

$$n - \sum_{i=1}^{k-1} n_i \text{ identical items of type } k.$$

- The number of visually distinguishable ways in which these n items can be arranged in a row is

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\cdots\binom{n-\sum_{i=1}^{k-2} n_i}{n_{k-1}}$$

$$= \frac{n!}{n_1!n_2!\cdots n_{k-1}!(n-\sum_{i=1}^{k-1} n_i)!}.$$

- The coefficients

$$\binom{n}{n_1, n_2, \cdots, n_{k-1}} = \frac{n!}{n_1!n_2!\cdots n_{k-1}!(n-\sum_{i=1}^{k-1} n_i)!}, \tag{1}$$

$$n_i \geq 0, i = 1, \ldots, k-1, \sum_{i=1}^{k-1} n_i \leq n$$

are called multinomial coefficients.

Remark 1.

- For $k = 2$ (so that $0 \leq n_1 \leq n$), the multinomial coefficients (1) reduce to binomial coefficients

$$\binom{n}{n_1} = \frac{n!}{n_1!(n-n_1)!}, n_1 \in \{0, 1, \ldots, n\}.$$

- For real numbers $x_1, \ldots, x_k$,

$$(x_1 + \cdots + x_k)^n = \underbrace{(x_1 + \cdots + x_k)(x_1 + \cdots + x_k)\ldots(x_1 + \cdots + x_k)}_{\text{product of n quantities}}.$$

A typical term in expansion of above product is an arrangement of $n_1\ x_1's, n_2\ x_2's, \ldots, n_{k-1}\ x_{k-1}'s$ and $(n-\sum_{i=1}^{k-1} n_i)\ x_k's$, $n_i \in \{0,1,\ldots\}$, $n_1+n_2+\cdots+n_{k-1} \le n$ (such as $x_1\ x_3\ x_4\ x_2\ x_1\ x_2 \cdots x_{k-2}\ x_8$). Each such term equals $x_1^{n_1}x_2^{n_2}\cdots x_k^{n_k}$ and total number of visually distinguishable ways of arranging

$$n_1\ x_1's, n_2\ x_2's, \ldots, n_{k-1}\ x_{k-1}'s \text{ and } \left(n-\sum_{i=1}^{k-1} n_i\right)x_k's \text{ is } \binom{n}{n_1,n_2,\cdots,n_{k-1}}.$$

Thus,

$$(x_1+\ldots+x_k)^n = \sum_{n_1=0}^{n}\cdots\sum_{\substack{n_{k-1}=0 \\ n_1+n_2+\cdots+n_{k-1}\le n}}^{n} \binom{n}{n_1,n_2,\cdots,n_{k-1}} x_1^{n_1}x_2^{n_2}\cdots x_{k-1}^{n_{k-1}}x_k^{n-\sum\limits_{i=1}^{k-1} n_i}.$$

Example 1. (Multinomial Distribution)

- A random experiment can result in one of $p+1$ $(p \geq 1)$ possible outcomes $A_1, A_2, \ldots, A_{p+1}$, where $A_i \cap A_j = \phi$, $i \neq j$ and $\cup_{i=1}^{p+1} A_i = \Omega$ (sample space). Let $P(A_i) = \theta_i \in (0,1)$, $i = 1, \ldots, p$, and $\sum_{i=1}^{p} \theta_i < 1$ so that $P(A_{p+1}) = 1 - \sum_{i=1}^{p} \theta_i \in (0,1)$. Suppose that the random experiment is repeated n times independently.

  Define

  $$X_i = \text{ number of times event } A_i \text{ occurs in } n \text{ trials}, \ i = 1, \ldots, p+1.$$

  Then one may be interested in the joint probability distribution of random variables $X_1, X_2, \ldots, X_{p+1}$.

Note that

$$X_{p+1} = n - \sum_{i=1}^{p} X_i = \text{ number of times } A_{p+1} \text{ occurs}$$

is completely determined by $\underline{X} = (X_1, X_2, \ldots, X_p)$ and, therefore, only distribution of $\underline{X} = (X_1, X_2, \ldots, X_p)$ may be of interest.

$$S_{\underline{X}} = \{\underline{x} = (x_1, x_2, \ldots, x_p) : x_i \in \{0, 1, \ldots, n\}, i = 1, \ldots, p, \sum_{i=1}^{p} x_i \leq n\}.$$

- The joint p.m.f. of $\underline{X} = (X_1, \ldots, X_p)$ is

$$f_{\underline{X}}(x_1, \ldots, x_p) = P(\{X_1 = x_1, \ldots, X_p = x_p\})$$

$$= \begin{cases} \frac{n!}{x_1! \cdots x_p! \left(n - \sum\limits_{i=1}^{p} x_i\right)!} \theta_1^{x_1} \ldots \theta_p^{x_p} \left(1 - \sum\limits_{i=1}^{p} \theta_i\right)^{(n - \sum_{i=1}^{p} x_i)}, & \text{if } \underline{x} \in S_{\underline{X}} \\ 0, & \text{otherwise} \end{cases}. \tag{2}$$

Definition 1.

The probability distribution given by (2) is called a multinomial distribution with n trials and cell probabilities $\theta_1, \ldots, \theta_p$ $\Big($denoted by $\text{Mult}(n, \theta_1, \ldots, \theta_p)$ $\Big)$.

- Note that,for $p = 1$, $\text{Mult}(n, \theta_1)$ distribution is nothing but the $\text{Bin}(n, \theta_1)$ distribution.

Result 1.

Let $\underline{X} = (X_1, X_2, \ldots, X_p) \sim \text{Mult}(n, \theta_1, \ldots, \theta_p)$, where $n \in \{1, 2, \ldots\}$, $\theta_i \in (0, 1)$, $i = 1, \ldots, p$, and $\sum_{i=1}^{p} \theta_i < 1$. Then

(i) $X_i \sim \text{Bin}(n, \theta_i)$, $i = 1, \ldots, p$;

(ii) $X_i + X_j \sim \text{Bin}(n, \theta_i + \theta_j)$, $i, j = 1, \ldots, p$, $i \neq j$;

(iii) $E(X_i) = n\theta_i$ and $\text{Var}(X_i) = n\theta_i(1 - \theta_i)$, $i = 1, \ldots, p$;

(iv) $\text{Cov}(X_i, X_j) = -n\theta_i\theta_j$, $i, j = 1, \ldots, p$, $i \neq j$.

Proof.

(i) Fix $i \in \{1, \ldots, p\}$. In a given trial of the random experiment, treat the occurrence of outcome A_i as success and that of any other A_j, $j \neq i$ (i.e., non-occurrence of A_i) as failure. Then we have a sequence of n independent Bernoulli trials with probability of success in each trial as $P(A_i) = \theta_i$. Therefore,

$$X_i = \text{Number of successes in } n \text{ independent Bernoulli trials } \sim \text{Bin}(n, \theta_i)$$

(ii) Fix $i, j \in \{1, \ldots, p\}$, $i \neq j$. In a given trial of the random experiment, treat the occurrence of outcome A_i or A_j (i.e., occurrence of $A_i \cup A_j$) as success and its non-occurrence as failure. Then we have a sequence of n independent Bernoulli trials with probability of success in each trial as $P(A_i \cup A_j) = P(A_i) + P(A_j) = \theta_i + \theta_j$ and, therefore,

$$\begin{aligned} X_i + X_j &= \text{Number of successes in } n \text{ independent Bernoulli trials} \\ &\sim \text{Bin}(n, \theta_i + \theta_j). \end{aligned}$$

(iii) Follows from (i) on using properties of Binomial distribution.

(iv) Fix $i, j \in \{1, \ldots, p\}$, $i \neq j$. Then

$$
\begin{aligned}
& X_i + X_j \sim \text{ Bin}(n, \theta_i + \theta_j) \\
\Rightarrow \quad & \text{Var}(X_i + X_j) = n(\theta_i + \theta_j)(1 - \theta_i - \theta_j) \\
\Rightarrow \quad & \text{Var}(X_i) + \text{Var}(X_j) + 2\text{Cov}(X_i, X_j) = n(\theta_i + \theta_j)(1 - \theta_i - \theta_j) \\
\Rightarrow \quad & n\theta_i(1 - \theta_i) + n\theta_j(1 - \theta_j) + 2\text{Cov}(X_i, X_j) \\
& \qquad\qquad\qquad\qquad = n(\theta_i + \theta_j)(1 - \theta_i - \theta_j) \\
\Rightarrow \quad & \text{Cov}(X_i, X_j) = -n\theta_i\theta_j, i \neq j.
\end{aligned}
$$

- The joint m.g.f. of $\underline{X} = (X_1, X_2, \ldots, X_p) \sim \text{Mult}(n, \theta_1, \ldots, \theta_p)$ is given by

$$M_{\underline{X}}(t) = \sum_{\substack{x_1=0 \\ x_1+\cdots+x_p \le n}}^{n} \cdots \sum_{x_p=0}^{n} e^{t_1 x_1 + \cdots + t_p x_p} \frac{n!}{x_1! \cdots x_p! \left(n - \sum\limits_{i=1}^{p} x_i\right)!} \theta_1^{x_1} \cdots \theta_p^{x_p} \left(1 - \sum_{i=1}^{p} \theta_i\right)^{n - \sum_{i=1}^{p} x_i}$$

$$= \sum_{\substack{x_1=0 \\ x_1+\cdots+x_p \le n}}^{n} \cdots \sum_{x_p=0}^{n} \frac{n!}{x_1! \cdots x_p! \left(n - \sum_{i=1}^{p} x_i\right)!} (\theta_1 e^{t_1})^{x_1} \cdots (\theta_p e^{t_p})^{x_p} \left(1 - \sum_{i=1}^{p} \theta_i\right)^{n - \sum_{i=1}^{p} x_i}$$

$$= \left(\theta_1 e^{t_1} + \cdots + \theta_p e^{t_p} + 1 - \sum_{i=1}^{p} \theta_i\right)^n, \ \underline{t} \in \mathbb{R}^p.$$

Bivariate Normal Distribution

Definition 2.

A bivariate random vector $\underline{X} = (X_1, X_2)$ is said to have a *bivariate normal distribtuion* $\mathrm{N}_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ if, for some $-\infty < \mu_i < \infty$, $i = 1, 2$, $\sigma_i > 0$, $i = 1, 2$, and $-1 < \rho < 1$, the joint p.d.f. of $\underline{X} = (X_1, X_2)$ is given by

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]},$$

$$\underline{x} = (x_1, x_2) \in \mathbb{R}^2.$$

- Note that $f_{X_1,X_2}(\underline{x}) \geq 0$, $\forall \underline{x} \in \mathbb{R}^2$ and on making the transformation $z_1 = \frac{x_1 - \mu_1}{\sigma_1}$ and $z_2 = \frac{x_2 - \mu_2}{\sigma_2}$ in the interval below, we have

$$
\begin{aligned}
I &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2) d\underline{x} \\
&= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}(z_1^2 - 2\rho z_1 z_2 + z_2^2)} d\underline{z} \\
&= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}(z_2^2 - \rho^2 z_2^2)} \underbrace{\left\{ \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}(z_1 - \rho z_2)^2} dz_1 \right\}}_{=\sqrt{1-\rho^2}\sqrt{2\pi}} dz_2
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z_2^2}{2}} dz_2 \\
&= 1.
\end{aligned}
$$

Therefore $f_{X_1,X_2}(x_1, x_2)$ is a p.d.f.

Result 2.

Suppose that $\underline{X} = (X_1, X_2) \sim \mathrm{N}_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, $-\infty < \mu_i < \infty$, $i = 1, 2$, $\sigma_i > 0$, $i = 1, 2$, and $-1 < \rho < 1$. Then,

(i) $X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$;

(ii) for a fixed $x_2 \in \mathbb{R}$, $X_1 | X_2 = x_2 \sim \mathrm{N}\Big(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2), \sigma_1^2(1 - \rho^2)\Big)$;

(iii) for a fixed $x_1 \in \mathbb{R}$, $X_2 | X_1 = x_1 \sim \mathrm{N}\Big(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1 - \mu_1), \sigma_2^2(1 - \rho^2)\Big)$;

(iv) the m.g.f. of $\underline{X} = (X_1, X_2)$ is

$$M_{X_1,X_2}(t_1, t_2) = e^{\mu_1 t_1 + \mu_2 t_2 + \frac{\sigma_1^2 t_1^2}{2} + \frac{\sigma_2^2 t_2^2}{2} + \rho\sigma_1\sigma_2 t_1 t_2}, \underline{t} = (t_1, t_2) \in \mathbb{R}^2;$$

(v) for real constants c_1 and c_2 such that $c_1^2 + c_2^2 > 0$

$$c_1X_1 + c_2X_2 \sim N(c_1\mu_1 + c_2\mu_2, c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + 2\rho c_1c_2\sigma_1\sigma_2);$$

(vi) $\rho(X_1, X_2) = \rho$;

(vii) X_1 and X_2 are independent if, and only if, $\rho = 0$.

Proof.

(i) For $x_1 \in \mathbb{R}$,

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1,X_2}(x_1, x_2)\,dx_2$$

$$= \frac{e^{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{x_2-\mu_2}{\sigma_2}-\rho\frac{x_1-\mu_1}{\sigma_1}\right]^2} dx_2$$

$$= \frac{e^{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma_2^2(1-\rho^2)}\left[x_2-\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1)\right)\right]^2} dx_2$$

$$= \frac{e^{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}}}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}$$

$$= \frac{1}{\sigma_1\sqrt{2\pi}}e^{-\frac{(x_1-\mu_1)^2}{2\sigma_1^2}},$$

which is the p.d.f. of $\mathrm{N}(\mu_1, \sigma_1^2)$ distribution. Thus $X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$. By symmetry $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$.

(ii) Fix $x_2 \in \mathbb{R}$,

$$\begin{aligned}
f_{X_1|X_2}(x_1|x_2) &= c_1(x_2) f_{X_1,X_2}(x_1, x_2) \\
&= c_2(x_2) e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}-\rho\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right)^2\right]} \\
&= c_2(x_2) e^{-\frac{1}{2\sigma_1^2(1-\rho^2)}\left(x_1-\left(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(x_2-\mu_2)\right)\right)^2}, x_1 \in \mathbb{R},
\end{aligned}$$

where $c_2(x_2)$ is the normalizing constant, i.e., $c_2(x_2)$ satisfies

$$\int_{-\infty}^{\infty} f_{X_1|X_2}(x_1|x_2) dx_1 = 1.$$

Clearly, for a fixed $x_2 \in \mathbb{R}$, $f_{X_1|X_2}(\cdot|x_2)$ is the p.d.f. of $\mathrm{N}\left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2 - \mu_2), \sigma_1^2(1 - \rho^2)\right)$ distribution.

(iii) Follows from (ii) on using symmetry.

(iv) For $\underline{t} = (t_1, t_2) \in \mathbb{R}^2$, we have

$$\begin{aligned} M_{X_1,X_2}(t_1, t_2) &= E(e^{t_1 X_1 + t_2 X_2}) \\ &= E\Big(E(e^{t_1 X_1 + t_2 X_2} | X_2)\Big) \\ &= E\Big(e^{t_2 X_2} E(e^{t_1 X_1} | X_2)\Big). \end{aligned}$$

For a fixed $x_2 \in \mathbb{R}$, since
$X_1 | X_2 = x_2 \sim \mathrm{N}\Big(\mu_1 + \rho \frac{\sigma_1}{\sigma_2}(x_2 - \mu_2), \sigma_1^2(1 - \rho^2)\Big)$, we have

$$E(e^{t_1 X_1} | X_2 = x_2) = e^{\left\{\mu_1 + \rho \frac{\sigma_1}{\sigma_2}(x_2 - \mu_2)\right\} t_1 + \frac{\sigma_1^2 (1-\rho^2) t_1^2}{2}}, \ t_1 \in \mathbb{R}.$$

Therefore, for $\underline{t} = (t_1, t_2) \in \mathbb{R}^2$,

$$M_{X_1,X_2}(t_1, t_2) = E\left(e^{t_2 X_2} e^{\left\{\mu_1+\rho\frac{\sigma_1}{\sigma_2}(X_2-\mu_2)\right\}t_1+\frac{\sigma_1^2(1-\rho^2)t_1^2}{2}}\right)$$
$$= e^{\mu_1 t_1+\frac{\sigma_1^2(1-\rho^2)t_1^2}{2}-\rho\frac{\sigma_1}{\sigma_2}\mu_2 t_1} E\left(e^{\left(t_2+\rho\frac{\sigma_1}{\sigma_2}t_1\right)X_2}\right).$$

Since $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$, we have

$$M_{X_1,X_2}(t_1, t_2) = e^{\mu_1 t_1+\frac{\sigma_1^2(1-\rho^2)t_1^2}{2}-\rho\frac{\sigma_1}{\sigma_2}\mu_2 t_1} e^{\left(t_2+\rho\frac{\sigma_1}{\sigma_2}t_1\right)\mu_2+\frac{\sigma_2^2\left(t_2+\rho\frac{\sigma_1}{\sigma_2}t_1\right)^2}{2}}$$
$$= e^{\mu_1 t_1+\mu_2 t_2+\frac{\sigma_1^2 t_1^2}{2}+\frac{\sigma_2^2 t_2^2}{2}+\rho\sigma_1\sigma_2 t_1 t_2}, \ \underline{t} = (t_1, t_2) \in \mathbb{R}^2.$$

(v) Let c_1 and c_2 be real constants such that $c_1^2 + c_2^2 > 0$ and let $Y = c_1X_1 + c_2X_2$. Then, for $t \in \mathbb{R}$,

$$\begin{aligned} M_Y(t) &= E(e^{tY}) \\ &= E(e^{tc_1X_1 + tc_2X_2}) \\ &= M_{X_1,X_2}(tc_1, tc_2) \\ &= e^{(c_1\mu_1 + c_2\mu_2)t + \frac{(c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + 2\rho c_1 c_2 \sigma_1\sigma_2)t^2}{2}}, \end{aligned}$$

which is the m.g.f. of $\mathrm{N}(c_1\mu_1 + c_2\mu_2, c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + 2\rho c_1 c_2 \sigma_1\sigma_2)$ distribution. Thus,

$$Y \sim \mathrm{N}(c_1\mu_1 + c_2\mu_2, c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + 2\rho c_1 c_2 \sigma_1\sigma_2).$$

(vi) By (i), $\text{Var}(X_1) = \sigma_1^2$ and $\text{Var}(X_2) = \sigma_2^2$. Also, for $\psi_{X_1,X_2}(t_1, t_2) = ln\,(M_{X_1,X_2}(t_1, t_2))$, $\underline{t} = (t_1, t_2) \in \mathbb{R}^2$,

$$\text{Cov}(X_1, X_2) = \left[\frac{\partial^2}{\partial t_1 \partial t_2}\psi_{X_1,X_2}(t_1, t_2)\right]_{\underline{t}=\underline{0}} = \rho\sigma_1\sigma_2$$

$$\Rightarrow \rho(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)\text{Var}(X_2)}} = \rho.$$

(vii) Since independent random variables are uncorrelated, it follows from (vi) that if X_1 and X_2 are independent then $\rho = 0$. Conversely, suppose that $\rho = 0$, Then, for $\underline{x} = (x_1, x_2) \in \mathbb{R}^2$,

$$f_{X_1,X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]}$$

$$= f_{X_1}(x_1) f_{X_2}(x_2),$$

implying that X_1 and X_2 are independent.

Result 3.

Let $\underline{X} = (X_1, X_2)$ be a bivariate random vector with $E(X_i) = \mu_i \in (-\infty, \infty)$, $\text{Var}(X_i) = \sigma_i^2 > 0$, $i = 1, 2$ and $\text{Cov}(X_1, X_2) = \rho \in (-1, 1)$. Then $\underline{X} \sim \text{N}_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ if, and only if, for any real constants t_1 and t_2 such that $t_1^2 + t_2^2 > 0$,

$$Y = t_1 X_1 + t_2 X_2 \sim \text{N}(t_1\mu_1 + t_2\mu_2, t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2\rho t_1 t_2 \sigma_1 \sigma_2).$$

Proof.

Clearly the necessary part of the assertion follows from Result 2 (v). Conversely, suppose that for all real constants t_1 and t_2 such that $t_1^2 + t_2^2 > 0$,

$$Y = t_1 X_1 + t_2 X_2 \sim \mathrm{N}(t_1\mu_1 + t_2\mu_2, t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2\rho t_1 t_2 \sigma_1\sigma_2).$$

Then, for $\underline{t} = (t_1, t_2) \in \mathbb{R}^2$,

$$\begin{aligned} M_{X_1,X_2}(t_1,t_2) &= E(e^{t_1 X_1 + t_2 X_2}) \\ &= E(e^Y) \\ &= M_Y(1) \\ &= e^{t_1\mu_1 + t_2\mu_2 + \frac{t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2\rho t_1 t_2 \sigma_1\sigma_2}{2}}, \end{aligned}$$

which is the m.g.f. of $\mathrm{N}_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Thus,

$$\underline{X} = (X_1, X_2) \sim \mathrm{N}_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho).$$

Take Home Problems

- Let $\underline{X} = (X_1, X_2, X_3)' \sim \text{Mult}(n, \theta_1, \theta_2, \theta_3)$. Find the conditional p.m.f. of (X_1, X_2) given $X_3 = x_3$, $x_3 \in \{0, 1, \ldots, n\}$;

- Let X_1 and X_2 be independent r.v.s with $X_i \sim \text{N}(\mu_i, \sigma_i^2)$, $i = 1, \ldots, k$. Find the distribution of $\underline{Y} = (X_1 + X_2, X_1 - X_2)$.

Thank you for your patience