IC105: Probability and Statistics **2021-22-M**

Lecture 20: Special Multivariate Distribution

Instructor: Dr. Kuldeep Kumar Kataria *Scribe:*

Example 20.1. *(a) Let X_1 and X_2 be independent r.v.'s with $X_i \sim GAM(\alpha_i, \theta)$, $\alpha_i > 0$, $\theta > 0$, $i = 1, 2$. Define $Y_1 = X_1 + X_2$ and $Y_2 = \frac{X_1}{X_1+X_2}$. Then Y_1 and Y_2 are independently distributed with $Y_1 \sim GAM(\alpha_1 + \alpha_2, \theta)$ and $Y_2 \sim Be(\alpha_1, \alpha_2)$.*

(b) Let X_1 and X_2 be iid $Exp(\theta)$ r.v.'s. Then $Y = \frac{X_1}{X_1+X_2} \sim U(0, 1)$.

Solutions: (a) The joint p.d.f. of $\underline{X} = (X_1, X_2)$ is

$$f_{\underline{X}}(x_1, x_2) = \prod_{i=1}^{2} f_{X_i}(x_i) = \prod_{i=1}^{2} \left\{ \frac{e^{-x_i/\theta} x_i^{\alpha_i - 1}}{\theta^{\alpha_i} \Gamma(\alpha_i)} I_{(0,\infty)}(x_i) \right\} = \begin{cases} \dfrac{e^{-(x_1+x_2)/\theta} x_1^{\alpha_1-1} x_2^{\alpha_2-1}}{\theta^{\alpha_1+\alpha_2}\Gamma(\alpha_1)\Gamma(\alpha_2)}, & \text{if } x_1 > 0,\ x_2 > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Here, $S_{\underline{X}} = (0, \infty)^2$. Let $h_1(X_1, X_2) = Y_1 = X_1 + X_2$ and $h_2(X_1, X_2) = Y_2 = \frac{X_1}{X_1+X_2}$. Thus $\underline{h} = (h_1, h_2) : S_{\underline{X}} \to \mathbb{R}^2$ is $1 - 1$ with inverse image (h_1^{-1}, h_2^{-1}), where

$$h_1^{-1}(y_1, y_2) = y_1 y_2,\ h_2^{-1}(y_1, y_2) = y_1(1 - y_2),\ J = \begin{vmatrix} y_2 & y_1 \\ 1 - y_2 & -y_1 \end{vmatrix} = -y_1.$$

$\underline{h}^{-1}(\underline{y}) \in S_{\underline{X}} \iff y_1 y_2 > 0,\ y_1(1 - y_2) > 0 \iff y_1 > 0,\ 0 < y_2 < 1 \implies h(S_{\underline{X}}) = (0, \infty) \times (0, 1)$. Thus the joint p.d.f. of $\underline{Y} = (Y_1, Y_2)$ is

$$\begin{aligned} f_{\underline{Y}}(y_1, y_2) &= \frac{e^{-(y_1 y_2 + y_1(1-y_2))/\theta} (y_1 y_2)^{\alpha_1-1} (y_1(1 - y_2))^{\alpha_2-1}}{\theta^{\alpha_1+\alpha_2}\Gamma(\alpha_1)\Gamma(\alpha_2)} | - y_1| I_{(0,\infty)\times(0,1)}(y_1, y_2) \\ &= \left\{ \frac{e^{-y_1/\theta} y_1^{\alpha_1+\alpha_2-1}}{\Gamma(\alpha_1 + \alpha_2)} I_{(0,\infty)}(y_1) \right\} \left\{ \frac{1}{B(\alpha_1, \alpha_2)} y_2^{\alpha_1-1} (1 - y_2)^{\alpha_2-1} I_{(0,1)}(y_2) \right\} = f_{Y_1}(y_1) f_{Y_2}(y_2), \end{aligned}$$

where $Y_1 \sim GAM(\alpha_1 + \alpha_2, \theta)$ and $Y_2 \sim Be(\alpha_1, \alpha_2)$. Clearly Y_1 and Y_2 are independent. Part (b) can similarly be proved.

20.1. Special Multivariate Distribution

20.1.1. Multinomial Distribution (A generalization of binomial distribution)

$\mathscr{E}$: a random experiment whose each trial results in one (and only one) of $p + 1$ possible outcomes $E_1, E_2, \ldots, E_{p+1}$ where $E_i \cap E_j = \phi$ and $\sum_{i=1}^{p+1} E_i = \Omega$. Let $P(A_i) = \theta_i \in (0, 1)$, $i = 1, 2 \ldots, p$ and $\sum_{i=1}^{p} \theta_i < 1$ so that $P(E_{p+1}) = 1 - \sum_{i=1}^{p} \theta_i \in (0, 1)$.

Consider n independent trials of $\mathscr{E}$. Define X_i = the number of times E_i occurs in n trials, $i = 1, 2 \ldots, p + 1$. Then $\sum_{i=1}^{p+1} X_i = n$, that is, $X_{p+1} = n - \sum_{i=1}^{n} X_i$. One may interested in probability distribution of $\underline{X} =$

$(X_1, X_2, \dots, X_p)$. We have

$$S_{\underline{X}} = \{\underline{x} = (x_1, x_2, \dots, x_p) : x_i \in \{0, 1, \dots, n\}, i = 1, 2, \dots, p, \ \sum_{i=1}^{n} x_i \le n\}$$

and

$$\begin{aligned} f_{\underline{X}}(x_1, x_2, \dots, x_p) &= P(X_1 = x_1, X_2 = x_2, \dots, X_p = x_p) \\ &= \begin{cases} \dfrac{n!}{x_1! x_2! \cdots x_p! (n - \sum_{i=1}^p x_i)!} \theta_1^{x_1} \theta_2^{x_2} \cdots \theta_p^{x_p} (1 - \sum_{i=1}^p \theta_i)^{n - \sum_{i=1}^p x_i}, & \underline{x} \in S_{\underline{X}}, \\ 0, \text{ otherwise.} \end{cases} \end{aligned}$$

$\longrightarrow$ Multinomial distribution with n trials and cell probabilities $\theta_1, \dots, \theta_p$ (denoted by $Mult(n, \theta_1, \theta_2, \dots, \theta_p)$)$\longrightarrow$ a family of distribution with varying $n \in \mathbb{N}$ and $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_p) \in \Theta = \{(t_1, t_2, \dots, t_p) : 0 < t_i < 1, \ i = 1, 2, \dots, p$ and $\sum_{i=1}^p t_i < 1\}$.

Remark 20.2. *For* $p = 1$, $Mult(n, \theta_1)$ *distribution is the same as* $Bin(n, \theta_1)$ *distribution.*

Theorem 20.3. *Suppose that* $\underline{X} = (X_1, X_2, \dots, X_p) \sim Mult(n, \theta_1, \theta_2, \dots, \theta_p)$, *where* $n \in \mathbb{N}$ *and* $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_p) \in \Theta$. *Then*

(a) $X_i \sim Bin(n, \theta_i)$, $i = 1, 2, \dots, p$,

(b) $X_i + X_j \sim Bin(n, \theta_i + \theta_j)$, $i = 1, 2, \dots, p$, $j = 1, 2, \dots, p$, $i \ne j$,

(c) $E(X_i) = n\theta_i$ *and* $\mathrm{Var}(X_i) = \sqrt{n\theta_i(1 - \theta_i)}$, $i = 1, 2, \dots, p$,

(d) $\mathrm{Cov}(X_i, X_j) = -n\theta_i\theta_j$, $i = 1, 2, \dots, p$, $j = 1, 2, \dots, p$, $i \ne j$.

Proof. (a) Fix $i \in \{1, 2, \dots, p\}$. A given trial of the experiment treat the occurrence of E_i as success and its non-occurrence (that is, occurrence of any other $E_j, j \ne i$) as failure. Then we have a sequence of independent Bernoulli trials with probability of success in each trial as $P(E_i) = \theta_i$. Thus

$$X_i = \text{the number of times } E_i \text{ occurs in } n \text{ Bernoulli trials} \sim Bin(n, \theta_i), i = 1, 2, \dots, p.$$

(b) Fix $i, j \in \{1, 2, \dots, p\}$ $i \ne j$. In any given trial of $\mathscr{E}$ consider occurrence of E_i or E_j as success and occurrence of any other E_l $(l \ne i, j)$ as failure. Then we have a sequence of n Bernoulli trials with success probability in each trials as $P(E_i + E_j) = \theta_i + \theta_j$,

$$X_i + X_j = \text{the number of success occurs in } n \text{ Bernoulli trials} \sim Bin(n, \theta_i + \theta_j).$$

(c) Obvious.

(d)

$$\begin{aligned} & \mathrm{Var}(X_i + X_j) = n(\theta_i + \theta_j)(1 - \theta_i - \theta_j) \\ \implies & \mathrm{Var}(X_i) + \mathrm{Var}(X_j) + 2\,\mathrm{Cov}(X_i, X_j) = n(\theta_i + \theta_j)(1 - \theta_i - \theta_j) \\ \implies & n\theta_i(1 - \theta_i) + n\theta_j(1 - \theta_i) + 2\,\mathrm{Cov}(X_i, X_j) = n(\theta_i + \theta_j)(1 - \theta_i - \theta_j) \implies \mathrm{Cov}(X_i, X_j) = -n\theta_i\theta_j. \end{aligned}$$

This completes the proof. $\square$

The m.g.f. of $\underline{X} = (X_1, X_2, \ldots, X_p)$ is given by

$$\begin{aligned}
M_{\underline{X}}(t_1, t_2, \ldots, t_p) &= E(e^{t_1X_1+t_2X_2+\cdots+t_pX_p}) \\
&= \sum_{\substack{x_1=0 \\ x_1+x_2+\cdots+x_p\le n}}^{n} \sum_{x_2=0}^{n} \cdots \sum_{x_p=0}^{n} e^{t_1x_1+t_2x_2+\cdots+t_px_p} \frac{n!\theta_1^{x_1}\theta_2^{x_2}\cdots\theta_p^{x_p}}{x_1!x_2!\cdots x_p!(n-\sum_{i=1}^{p} x_i)!}(1-\sum_{i=1}^{p}\theta_i)^{n-\sum_{i=1}^{p} x_i} \\
&= \sum_{\substack{x_1=0 \\ x_1+x_2+\cdots+x_p\le n}}^{n} \sum_{x_2=0}^{n} \cdots \sum_{x_p=0}^{n} e^{t_1x_1+t_2x_2+\cdots+t_px_p} \frac{n!(\theta_1e^{t_1})^{x_1}(\theta_2e^{t_2})^{x_2}\cdots(\theta_pe^{t_p})^{x_p}}{x_1!x_2!\cdots x_p!(n-\sum_{i=1}^{p} x_i)!}(1-\sum_{i=1}^{p}\theta_i)^{n-\sum_{i=1}^{p} x_i} \\
&= \left(\theta_1e^{t_1}+\cdots+\theta_pe^{t_p}+1-\sum_{i=1}^{p}\theta_i\right)^n, \ \underline{t}\in\mathbb{R}^p.
\end{aligned}$$

Remark 20.4. *The last theorem can also be proved using m.g.f. For example (for $i, j \in \{1, 2, \ldots, p\}, i \neq j$)*

$$M_{X_i+X_j}(\underline{t}) = M_{\underline{X}}(0,\ldots,0,t_i,0,\ldots,0,t_j,0,\ldots,0) = ((\theta_i+\theta_j)e^t+1-\theta_i-\theta_j), \ \ \underline{t}\in\mathbb{R}^p.$$

20.1.2. Bivariate Normal Distribution

Definition 20.5. *A bivariate r.v. $\underline{X} = (X_1, X_2)$ is said to follow bivariate normal distribution $N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ if for some $-\infty < \mu_i < \infty$, $\sigma_i > 0$, $i = 1, 2$ and $-1 < \rho < 1$, the joint p.d.f. of $\underline{X} = (X_1, X_2)$ is given by*

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

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

$$\begin{aligned}
\int_{\mathbb{R}^2} f_{X_1,X_2}(x_1,x_2)\mathrm{d}x_1\mathrm{d}x_2 &= \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}(z_1^2-2\rho z_1z_2+z_2^2)}\mathrm{d}z_1\mathrm{d}z_2 \\
&= \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{-\infty}^{\infty} e^{-\frac{z_2^2}{2}}\left\{\int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}(z_1-\rho z_2)^2}\mathrm{d}z_2\right\}\mathrm{d}z_1 \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{z_2^2}{2}}\mathrm{d}z_2 = 1 \implies f_{X_1,X_2}(x_1,x_2) \text{ is a p.d.f.}
\end{aligned}$$

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

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

$$\implies X_1|X_2 = x_2 \sim N\left(\mu_1+\frac{\rho\sigma_1}{\sigma_2}(x_2-\mu_2), \sigma_1^2(1-\rho^2)\right), \ X_2 \sim N(\mu_2, \sigma_2^2).$$

By symmetry

$$X_2|X_1 = x_1 \sim N\left(\mu_2+\frac{\rho\sigma_2}{\sigma_1}(x_1-\mu_1), \sigma_2^2(1-\rho^2)\right), \ X_1 \sim N(\mu_1, \sigma_1^2).$$

Clearly, $\mu_1 = E(X_1),\ \mu_2 = E(X_2),\ \sigma_1^2 = \mathrm{Var}(X_1)$ and $\sigma_2^2 = \mathrm{Var}(X_2)$.

$$\begin{aligned}
\text{m.g.f. } M_{X_1,X_2}(t_1,t_2) &= E\left(e^{t_1X_1+t_2X_2}\right)\\
&= E\left(E\left(e^{t_1X_1+t_2X_2}|X_2\right)\right) = E\left(e^{t_2X_2}E\left(e^{t_1X_1}|X_2\right)\right),\ \underline{t}=(t_1,t_2)\in\mathbb{R}^2,\\
E\left(e^{t_1X_1}|X_2\right) &= \text{m.g.f. of conditional distribution } X_1|X_2 \text{ at point } t_2\\
&= e^{\{\mu_1+\frac{\rho\sigma_1}{\sigma_2}(X_2-\mu_2)\}t_1+\frac{\sigma_1^2(1-\rho^2)t_1^2}{2}}\\
M_{X_1,X_2}(t_1,t_2) &= e^{\{\mu_1-\frac{\rho\sigma_1}{\sigma_2}\mu_2\}t_1+\frac{\sigma_1^2(1-\rho^2)t_1^2}{2}}E\left[e^{t_2X_2}e^{\frac{\rho\sigma_1}{\sigma_2}t_1X_2}\right]\\
&= e^{\{\mu_1-\frac{\rho\sigma_1}{\sigma_2}\mu_2\}t_1+\frac{\sigma_1^2(1-\rho^2)t_1^2}{2}}M_{X_2}\left[t_2+\frac{\rho\sigma_1}{\sigma_2}t_1\right]\\
&= e^{\{\mu_1-\frac{\rho\sigma_1}{\sigma_2}\mu_2\}t_1+\frac{\sigma_1^2(1-\rho^2)t_1^2}{2}}e^{\mu_2\{t_2+\frac{\rho\sigma_1}{\sigma_2}t_1\}+\frac{\sigma_2^2}{2}(t_2+\frac{\rho\sigma_1}{\sigma_2}t_1)^2}\\
&= e^{\mu_1t_1+\mu_2t_2+\frac{\sigma_1^2t_1^2}{2}+\frac{\sigma_2^2t_2^2}{2}+\rho\sigma_1\sigma_2t_1t_2},\ \underline{t}=(t_1,t_2)\in\mathbb{R}^2.
\end{aligned}$$

Thus we have the following theorem.

Theorem 20.6. *Suppose that $\underline{X} = (X_1,X_2) \sim N_2(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$, $-\infty < \mu_i < \infty$, $\sigma_i > 0$, $i = 1,2$ and $-1 < \rho < 1$. Then*

(a) $X_1 \sim N(\mu_1,\sigma_1^2)$ and $X_2 \sim N(\mu_2,\sigma_2^2)$;

(b) For fixed $x_2 \in \mathbb{R}$, $X_1|X_2 = x_2 \sim N\left(\mu_1 + \frac{\rho\sigma_1}{\sigma_2}(x_2-\mu_2), \sigma_1^2(1-\rho^2)\right)$ and for fixed $x_1 \in \mathbb{R}$, $X_2|X_1 = x_1 \sim N\left(\mu_2 + \frac{\rho\sigma_2}{\sigma_1}(x_1-\mu_1), \sigma_2^2(1-\rho^2)\right)$;

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

$$M_{X_1,X_2}(t_1,t_2) = e^{\mu_1t_1+\mu_2t_2+\frac{\sigma_1^2t_1^2}{2}+\frac{\sigma_2^2t_2^2}{2}+\rho\sigma_1\sigma_2t_1t_2},\ \underline{t}=(t_1,t_2)\in\mathbb{R}^2;$$

(d) $\rho(X_1,X_2) = \mathrm{Corr}(X_1,X_2) = \rho$;

(e) X_1 and X_2 are independent iff $\rho = 0$;

(f) For real constants C_1 and C_2 such that $(C_1,C_2) \neq (0,0)$

$$C_1X_1 + C_2X_2 \sim N(C_1\mu_1 + C_2\mu_2, C_1^2\sigma_1^2 + C_2^2\sigma_2^2 + 2\rho C_1C_2\sigma_1\sigma_2)$$

Proof. (a)-(c) Already done.

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

$$\begin{aligned}
\psi_{X_1,X_2}(t_1,t_2) &= \ln M_{X_1,X_2}(t_1,t_2) = \mu_1t_1+\mu_2t_2+\frac{\sigma_1^2t_1^2}{2}+\frac{\sigma_2^2t_2^2}{2}+\rho\sigma_1\sigma_2t_1t_2\\
\frac{\partial}{\partial t_1}\psi_{X_1,X_2}(t_1,t_2) &= \mu_1+2\sigma_1^2t_1+\rho\sigma_1\sigma_2t_2\\
\frac{\partial^2}{\partial t_2\partial t_1}\psi_{X_1,X_2}(t_1,t_2) &= \rho\sigma_1\sigma_2\\
\implies \mathrm{Cov}(X_1,X_2) &= \left[\frac{\partial^2}{\partial t_2\partial t_1}\psi_{X_1,X_2}(t_1,t_2)\right]_{\underline{t}=\underline{0}} = \rho\sigma_1\sigma_2\\
\implies \rho(X_1,X_2) = \mathrm{Corr}(X_1,X_2) &= \frac{\mathrm{Cov}(X_1,X_2)}{\sqrt{\mathrm{Var}(X_1)\,\mathrm{Var}(X_2)}} = \rho.
\end{aligned}$$

(e) Obviously, if X_1 and X_2 are independent then $\rho = \mathrm{Corr}(X_1, X_2) = 0$. Now suppose that $\rho = 0$. Then

$$\begin{aligned} f_{X_1,X_2}(x_1,x_2) &= \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]} \\ &= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{1}{2\sigma_1^2}[x_1-\mu_1]^2} \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2\sigma_2^2}(x_2-\mu_2)^2} \\ &= f_{X_1}(x_1) f_{X_2}(x_2)\ \forall\, \underline{x} = (x_1,x_2) \in \mathbb{R}^2 \implies X_1 \text{ and } X_2 \text{ are independent.} \end{aligned}$$

(f) Let $Y = C_1X_1 + C_2X_2$. Then

$$\begin{aligned} M_Y(t) &= E(e^{tY}) = E(e^{t(C_1X_1+C_2X_2)}) = M_{X_1,X_2}(tC_1, tC_2) \\ &= \exp\left\{C_1t\mu_1 + C_2t\mu_2 + \frac{C_1^2t^2\sigma_1^2}{2} + \frac{C_2^2t^2\sigma_2^2}{2} + \rho t^2 C_1C_2\sigma_1\sigma_2\right\} \\ &= \exp\left\{(C_1\mu_1 + C_2\mu_2)t + (C_1^2\sigma_1^2 + C_2^2\sigma_2^2 + 2\rho C_1C_2\sigma_1\sigma_2)\frac{t^2}{2}\right\} \\ &\longrightarrow \text{m.g.f. of } N(C_1\mu_1 + C_2\mu_2, C_1^2\sigma_1^2 + C_2^2\sigma_2^2 + 2\rho C_1C_2\sigma_1\sigma_2). \end{aligned}$$

This completes the proof. □

Theorem 20.7. *Let $\underline{X} = (X_1, X_2)$ be a bivariate r.v. with $E(X_i) = \mu_i \in (-\infty, \infty)$, $\mathrm{Var}(X_i) = \sigma_i^2$, ($\sigma_i > 0$), $i = 1, 2$ and $\mathrm{Corr}(X_1, X_2) = \rho \in (-1, 1)$. Then $\underline{X} \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ iff for any $\underline{t} = (t_1, t_2) \in \mathbb{R}^2 - \{0\}$, $Y = t_1X_1 + t_2X_2 \sim N(t_1\mu_1 + t_2\mu_2, t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2\rho t_1t_2\sigma_1\sigma_2)$.*

Proof. Let $\underline{X} \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Then by (f) of last theorem

$$Y = t_1X_1 + t_2X_2 \sim N(t_1\mu_1 + t_2\mu_2, t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2\rho t_1t_2\sigma_1\sigma_2)\ \forall\, \underline{t} \in \mathbb{R}^2 - \{0\}.$$

Conversely, suppose that for all $\underline{t} = (t_1, t_2) \in \mathbb{R}^2 - \{0\}$, $Y = t_1X_1 + t_2X_2 \sim N(t_1\mu_1 + t_2\mu_2, t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2\rho t_1t_2\sigma_1\sigma_2)$. Then for $\underline{t} = (t_1, t_2) \in \mathbb{R}^2 - \{0\}$

$$\begin{aligned} M_{X_1,X_2}(t_1,t_2) &= E(e^{t_1X_1+t_2X_2}) \\ &= M_Y(1) = \exp\left\{(t_1\mu_1 + t_2\mu_2) + (t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2\rho t_1t_2\sigma_1\sigma_2)\frac{1}{2}\right\} \\ &\longrightarrow \text{m.g.f. of } N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) \implies \underline{X} = (X_1, X_2) \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho). \end{aligned}$$

This completes the proof. □

Theorem 20.8. *Let $X_1, X_2, \ldots, X_n$ ($n \geq 2$) be a random sample from $N(\mu, \sigma^2)$ distribution, where $-\infty < \mu < \infty$ and $\sigma > 0$. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$ denote the sample mean and sample variance, respectively. Then*

(i) $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$;

(ii) $\bar{X}$ *and* S^2 *are independent r.v.'s;*

(iii) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$;

(iv) $E(S^2) = \sigma^2$, $\mathrm{Var}(S^2) = \frac{2\sigma^4}{n-1}$, $E(S) = \sqrt{\frac{2}{n-1}}\frac{\Gamma(n/2)}{\Gamma(\frac{n-1}{2})}\sigma$.

Proof. (i) Follows from last theorem by taking $k = n$, $a_i = \frac{1}{n}$, $\mu_i = \mu$, $\sigma_i^2 = \sigma$, $i = 1, 2, \ldots, n$.

(ii) Let $Y_i = X_i - \bar{X}$, $i = 1, 2, \ldots, n$ and let $\underline{Y} = (Y_1, Y_2, \ldots, Y_n)$. Then

$$\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} (X_i - \bar{X}) = \sum_{i=1}^{n} X_i - n\bar{X} = 0$$

$$(n-1)S^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 = \sum_{i=1}^{n} Y_i^2 \quad \text{(a function of } \underline{Y}\text{)}$$

The joint m.g.f. of $(\underline{Y}, \bar{X})$ is given by

$$M_{\underline{Y},\bar{X}}(\underline{t}) = E\left(e^{\sum_{i=1}^{n} t_i Y_i + t_{n+1}\bar{X}}\right), \;\; \underline{t} = (t_1, t_2, \ldots, t_n, t_{n+1}) \in \mathbb{R}^{n+1}$$

$$\begin{aligned}
\sum_{i=1}^{n} t_i Y_i + t_{n+1}\bar{X} &= \sum_{i=1}^{n} t_i (X_i - \bar{X}) + t_{n+1}\bar{X} \\
&= \sum_{i=1}^{n} t_i X_i + \frac{(t_{n+1} - \sum_{i=1}^{n} t_i)}{n} \sum_{i=1}^{n} X_i \\
&= \sum_{i=1}^{n} \left(t_i - \bar{t} + \frac{t_{n+1}}{n}\right) X_i, \;\text{ where } \bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i \\
&= \sum_{i=1}^{n} u_i X_i, \;\text{ where } u_i = t_i - \bar{t} + \frac{t_{n+1}}{n}, \; i = 1(1)n.
\end{aligned}$$

Then $\sum_{i=1}^{n} u_i = t_{n+1}$ and $\sum_{i=1}^{n} u_i^2 = \sum_{i=1}^{n} (t_i - \bar{t})^2 + \frac{t_{n+1}^2}{n}$.

$$\begin{aligned}
M_{\underline{Y},\bar{X}}(\underline{t}) &= E\left(e^{\sum_{i=1}^{n} u_i X_i}\right) \\
&= \prod_{i=1}^{n} M_{X_i}(u_i) \\
&= \prod_{i=1}^{n} e^{\mu u_i + \frac{1}{2}\sigma^2 u_i^2} \\
&= e^{\mu \sum_{i=1}^{n} u_i + \frac{\sigma^2}{2}\sum_{i=1}^{n} u_i^2} \\
&= e^{\mu t_{n+1} + \frac{\sigma^2}{2}\{\sum_{i=1}^{n}(t_i - \bar{t})^2 + \frac{t_{n+1}^2}{n}\}} = \left\{e^{\mu t_{n+1} + \frac{\sigma^2 t_{n+1}^2}{2n}}\right\}\left\{e^{\frac{\sigma^2}{2}\sum_{i=1}^{n}(t_i - \bar{t})^2}\right\}
\end{aligned}$$

$$\begin{aligned}
& M_{\underline{Y}}(t_1, t_2, \ldots, t_n) = M_{\underline{Y},\bar{X}}(t_1, t_2, \ldots, t_n, 0) = e^{\frac{\sigma^2}{2}\sum_{i=1}^{n}(t_i - \bar{t})^2}, (t_1, t_2, \ldots, t_n) \in \mathbb{R}^n \\
& M_{\bar{X}}(t_{n+1}) = M_{\underline{Y},\bar{X}}(0, \ldots, 0, t_{n+1}) = e^{\mu t_{n+1} + \frac{\sigma^2 t_{n+1}^2}{2n}}, \; t_{n+1} \in \mathbb{R} \\
\Longrightarrow\; & M_{\underline{Y},\bar{X}}(\underline{t}) = M_{\underline{Y}}(t_1, t_2, \ldots, t_n) M_{\bar{X}}(t_{n+1}), \forall\, \underline{t} = (t_1, t_2, \ldots, t_n, t_{n+1}) \in \mathbb{R}^{n+1} \\
\Longrightarrow\; & \underline{Y} \text{ and } \bar{X} \text{are independent} \\
\Longrightarrow\; & \sum_{i=1}^{n}(X_i - \bar{X})^2 \text{ and } \bar{X} \text{ are independent.}
\end{aligned}$$

(iii) Let $Z_i = \frac{X_i - \mu}{\sigma}$, $i = 1, 2, \ldots, n$. Then $Z_1, Z_2, \ldots, Z_n$ are iid $N(0, 1)$ r.v.'s. Also let $Z = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0, 1)$ (using (i)). Let

$$W = Z^2 = \frac{n(\bar{X} - \mu)^2}{\sigma^2} \;\text{ and }\; T = \frac{(n-1)S^2}{\sigma^2}.$$

Then by (ii), W and T are independent r.v.s. Also $W \sim \chi^2_1$ and $V = \sum_{i=1}^n Z_i^2 \sim \chi^2_n$.

$$\begin{aligned}
V = \sum_{i=1}^n Z_i^2 &= \frac{1}{\sigma^2}\sum_{i=1}^n (X_i-\mu)^2 \\
&= \frac{1}{\sigma^2}\sum_{i=1}^n (X_i-\bar{X}+\bar{X}-\mu)^2 \\
&= \frac{1}{\sigma^2}\sum_{i=1}^n (X_i-\bar{X})^2 + \frac{n(\bar{X}-\mu)^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} + \frac{n(\bar{X}-\mu)^2}{\sigma^2} = T + W.
\end{aligned}$$

This implies

$$\begin{aligned}
&M_V(t) = M_T(t)M_W(t) \\
\implies & M_T(t) = \frac{M_V(t)}{M_W(t)} = \frac{(1-2t)^{-n/2}}{(1-2t)^{-1/2}} = (1-2t)^{-(n-1)/2}, \ \ t < \frac{1}{2} \to \text{m.g.f. of } \chi^2_{n-1} \\
\implies & T = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.
\end{aligned}$$

(iv) $T = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_v$, where $v = n-1$. Thus

$$\begin{aligned}
E(T^s) &= \int_0^\infty t^s \frac{1}{2^{v/2}\Gamma(v/2)} e^{-t/2} t^{v/2-1} \mathrm{d}t \\
&= \frac{1}{2^{v/2}\Gamma(v/2)} \int_0^\infty e^{-t/2} t^{\frac{v+2s}{2}-1} \mathrm{d}t = \frac{2^{\frac{v+2s}{2}}\Gamma\left(\frac{v+2s}{2}\right)}{2^{v/2}\Gamma(v/2)} = \frac{2^s\Gamma\left(\frac{v+2s}{2}\right)}{\Gamma(v/2)}, \ s > -\frac{v}{2}.
\end{aligned}$$

This implies

$$\begin{aligned}
\frac{(n-1)^s}{\sigma^{2s}} E(S^{2s}) &= \frac{2^s\Gamma\left(\frac{v+2s}{2}\right)}{\Gamma(v/2)} \\
\implies E(S^r) &= \left(\frac{2}{n-1}\right)^{r/2} \frac{\Gamma\left(\frac{v+r}{2}\right)}{\Gamma(v/2)} \sigma^r, \ \ r > v \\
\implies E(S^r) &= \left(\frac{2}{n-1}\right)^{r/2} \frac{\Gamma\left(\frac{n-1+r}{2}\right)}{\Gamma((n-1)/2)} \sigma^r, \ \ r > v \\
\implies E(S) &= \left(\frac{2}{n-1}\right)^{1/2} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma((n-1)/2)} \sigma \\
\implies E(S^2) &= \left(\frac{2}{n-1}\right) \frac{\Gamma\left(\frac{n-1}{2}+1\right)}{\Gamma((n-1)/2)} \sigma^2 = \sigma^2 \\
\implies E(S^4) &= \left(\frac{2}{n-1}\right)^2 \frac{\Gamma\left(\frac{n-1}{2}+2\right)}{\Gamma((n-1)/2)} \sigma^4 = \frac{n+1}{n-1}\sigma^4 \\
\operatorname{Var}(S^2) &= E(S^4) - (E(S^2))^2 = \frac{2\sigma^4}{n-1}.
\end{aligned}$$

This completes the proof. □

Remark 20.9. *Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution having p.m.f. / p.d.f. f. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i-\bar{X})^2$. Let $E(X_1) = \mu$ and $\operatorname{Var}(X_1) = \sigma^2$. Then $E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^n X_i\right) =$*

μ, $\text{Var}(\bar{X}) = \text{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{\sigma^2}{n}$.

$$\begin{aligned}
E\left[(n-1)S^2\right] &= E\left[\sum_{i=1}^{n}(X_i - \bar{X})^2\right] \\
\implies (n-1)E(S^2) &= E\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right] \\
&= \sum_{i=1}^{n} E(X_i^2) - nE(\bar{X}^2) \\
&= n\left[E(X_1^2) - E(\bar{X}^2)\right] \\
&= n\left[\text{Var}(X_1) + (E(X_1))^2 - \text{Var}(\bar{X}) - (E(\bar{X}))^2\right] \\
&= n(\sigma^2 + \mu^2 - \frac{\sigma^2}{n} - \mu^2) = (n-1)\sigma^2 \implies E(S^2) = \sigma^2.
\end{aligned}$$

For this reason $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ *is called sample variance and not* $S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$*. Note that* $E(S_1^2) = E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right) = \frac{n-1}{n}\sigma^2 < \sigma^2$*, i.e.,* S_1^2 *underestimates* σ^2*.*