Module 5

Some Special Distributions

5.1. Summary of Probability Distributions

Let X be a r.v. defined on a probability space $(\Omega, \mathcal{B}, P)$ associated with a random $\mathcal{E}$.

F_X: distribution function of X

f_X: p.m.f. / p.d.f. of X (i.e. p.d.f/p.m.f.)

The probability distribution of X describes the manner in which the rv X takes values in various sets. It may be desirable to have a set of numerical measures that provide a summary of the prominent features of the probability distribution of X. We call these measures as descriptive measures. Four prominently used descriptive measures are:

(1) Measures of central tendency or location (also called averages)

→ Give us the idea about central value of the probability distribution around which the value of rv X are clustered. Commonly used measures of central tendency are:

(a) Mean:

$$\mu_1' = E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx \quad \text{or} \quad \sum_{x \in S_X} x f_X(x)$$

→ may or may not exist

Whenever it exists it gives us the idea about average observed value of X when $\mathcal{E}$ is repeated a large number of times.

Note that if distribution of X is symmetric about μ (i.e., $X-\mu \overset{d}{=} \mu-X$) then

$$E(X)=\mu,$$

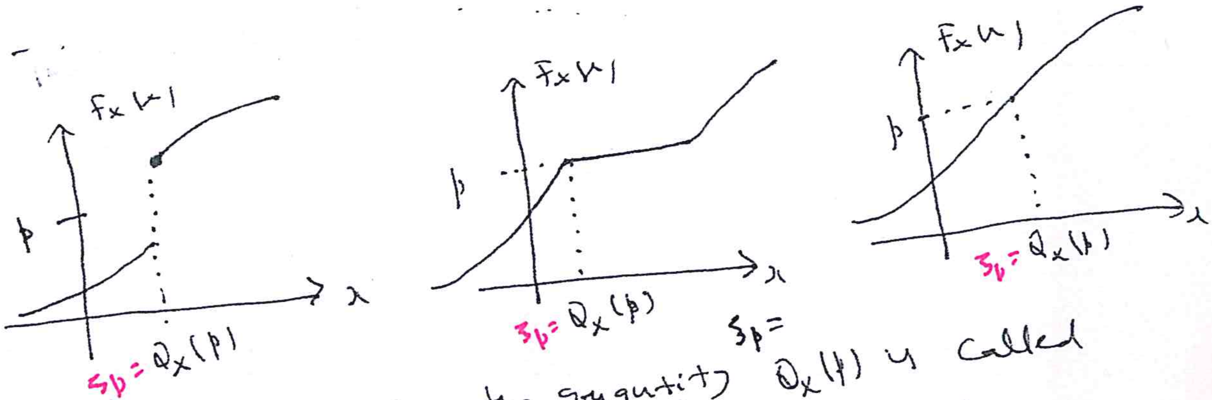
provided it exists.

Mean seems to be the best suited measure of central tendency for symmetric distribution. Because of its simplicity mean is the most commonly used average. However mean may be affected by a few extreme values and also it may not be defined.

$\boxed{\text{(b) Median}}$ Before defining the median we first introduce the concept of quantile function or quantile.

The quantile function of rv X is a function $Q_X : (0,1) \to \mathbb{R}$ defined by

$$Q_X(p) = \inf\{x \in \mathbb{R} : F_X(x) \geq p\} \quad p \in (0,1).$$



For a fixed $p \in (0,1)$ the quantity $\xi_p = Q_X(p)$ is called the quantile of order p. Note that

$$F_X(\xi_p-) \leq p \leq F_X(\xi_p), \quad \text{(Exercise)}$$

and $F_X(\xi_p) = p$ provided F_X is continuous at ξ_p

Also note that:

- $Q_X(F_X(x)) \leq x$, provided $0 < F_X(x) < 1$

- $F_X(Q_X(p)) \geq p, \quad \forall\, 0 < p < 1;$
- F_X is continuous $\Rightarrow F_X(Q_X(p)) = p;$
- $Q_X(p) \leq x \Leftrightarrow F_X(x) \geq p;$
- $Q_X(p) = F_X^{-1}(p)$ provided $F_X^{-1}(p)$ exists;
- $Q_X(p_1) \leq Q_X(p_2) \quad \forall\, 0 < p_1 < p_2 < 1$

The quantile of order 0.5 is called the median of (distribution) of X. If me is the median of X then

$$F_X(m_e-) \leq \frac{1}{2} \leq F_X(m_e)$$

If the random experiment $\mathcal{E}$ is repeated a large number of times about half of the times observed value of X is expected to be less than me and about half of the times it is expected to be greater than me

Suppose that the distribution of X is symmetric about μ. Then

$$X - \mu \overset{d}{=} \mu - X$$

$$\Rightarrow P(X - \mu \leq 0) = P(\mu - X \leq 0)$$

$$\Rightarrow F_X(\mu) = 1 - F_X(\mu-)$$

$$\Rightarrow F_X(\mu-) \leq \frac{1}{2} \leq F_X(\mu)$$

$\Rightarrow \mu = E(X) = m_e$, provided F_X is continuous at μ.

$\boxed{\text{Merits of median as a measure of central tendency:}}$

- Unlike mean it is always defined;
- Median is not affected by a few extreme values of X as it takes into account only the probabilities with which different values occur and not their numerical values;

As a measure of central tendency the median is preferred over the mean if the distribution is asymmetric and a few extreme observations occur with positive probability.

Demerits of Median as a measure of central tendency

- Does not at all take into account the numerical values assumed by X;
- For many probability distributions it is not easy to evaluate.

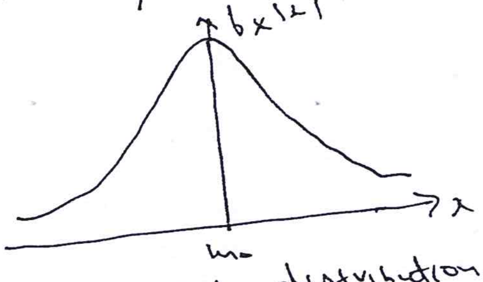
(c) Mode

Roughly speaking mode m_0 of a probability distribution is the value that occurs with highest probability and is defined by

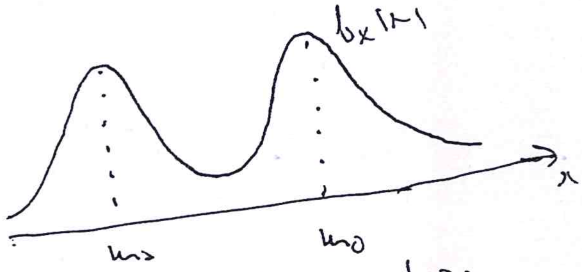
$$f_X(m_0) = \sup\{f_X(x) : x \in S_X\}$$

If the random experiment $\mathcal{E}$ is repeated a large number of times then either mode m_0 or a value in the neighborhood of m_0 is observed with maximum frequency.

Note that mode of a distribution may not be unique. A distribution having single/double/triple/multiple mode(s) is called a unimodal/bimodal/trimodal/multimodal distribution.



Unimodal distribution



Bimodal distribution

Trimodal distribution

Merits of a mode as a measure of central tendency

- It is easy to understand and easy to calculate. Normally, it can be found by just inspection

Demerits of mode as a measure of central tendency

- A probability distribution may have more than one mode which may be far apart

As a measure of central tendency mode is less preferred than mean and median. Clearly for symmetric unimodal distributions mean = median = mode

(ii) Measures of Dispersion

Apart from measures of central tendency other measures are often required to describe a probability distribution. The measures of dispersion give the idea about the scatter (cluster/dispersion) of probability mass of the distribution about a measure of central tendency. Some of the measures of dispersion are listed below.

(a) Range

Let $S_X = [a, b]$. Then range of distribution of X is defined by

$$R = b - a$$

It does not take into account how the probability mass is distributed over $[a, b]$. For this reason it is not a preferred measure of dispersion

(b) Mean Deviation

Let

A : suitable measure of central tendency.

Define

$MD(A) = E(|X-A|)$, provided it exists

↳ called mean deviation of X about A

$MD(\mu) = E(|X-\mu|)$, where $\mu = E(X)$

↳ mean deviation about mean μ

$MD(me) = E(|X-me|)$

↳ mean deviation about median

It can be shown that

$$MD(me) \leq MD(A) \quad \forall A \in \mathbb{R}$$

For this reason $MD(me)$ seems to be more appealing than $MD(A)$ for any $A \in \mathbb{R}$.

- $MD(A)$ is generally difficult to compute for many distributions.
- $MD(A)$ is sensitive to extreme observations.
- $MD(A)$ may not exist for many distributions.

(c) Standard Deviation

The standard deviation of distribution of X is defined by

$$\sigma = \sqrt{Var(X)} = \sqrt{E((X-\mu)^2)}, \quad \text{where } \mu \in \mathbb{R}$$

↳ has same unit as that of X.

Clearly

$$\sigma \leq \sqrt{E((X-A)^2)}, \quad \forall A \in \mathbb{R}$$

Standard deviation σ gives us the idea of average spread of values of X around mean μ.

- σ is simple to compute for most distributions (unlike $MD(A)$, $A \in \mathbb{R}$)

SD is
- most widely measure of dispersion (especially for nearly symmetric distributions)
- For some distributions SD does not exist
- SD is sensitive to extreme observations

(d) Quartile Deviation

Let

$q_1 = \xi_{0.25}$ = quantile of order 0.25 (lower quartile of X)

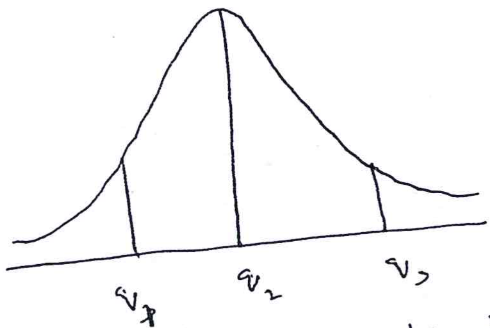
$q_2 = me = \xi_{0.5}$ = quantile of order 0.5 = median

and $q_3 = \xi_{0.75}$ = quartile of order 0.75 (upper quartile of X)

q_1, q_2, q_3 divide the probability distribution of X into 4 parts so that

$F_X(q_1-) \le \frac{1}{4} \le F_X(q_1)$; $F_X(q_2-) \le \frac{1}{2} \le F_X(q_2)$

and $F_X(q_3-) \le \frac{3}{4} \le F_X(q_3)$



Note that q_1 q_2 and q_3 divide the p.d.f. / p.m.f. of X into 4 parts so that each of them has 25% probability mass. Define

$IQR = q_3 - q_1 \rightarrow$ inter-quartile range

$QD = \dfrac{q_3 - q_1}{2} \rightarrow$ quartile deviation or the semi-interquartile range

- Unlike SD, QD is not sensitive to extreme values assumed by X.
- Does not at all take into account numerical values of X
- Ignores the tail of the probability distribution (constituting 50% of probability distribution on left side of Q_1 and right side of Q_3)
- QD depends on the unit of measurements of X and thus it may not be appropriate for comparing dispersions of two probability distributions having different units of measurements. For this purpose one may use

$$CQD = \frac{Q_3 - Q_1}{Q_3 + Q_1} \rightarrow \text{Coefficient of quartile deviation}$$

$\hookrightarrow$ does not depend on units of measurements.

(d) Coefficient of Variation

Like QD, the SD σ also depends on units of measurements of rv X and thus it is not an appropriate measure of dispersion for comparing distributions having different units of measurements. For this purpose we consider

$$CV = \frac{\sigma}{\mu} \rightarrow \text{Coefficient of variation}$$

$\hookrightarrow$ does not depend on units of measurements

where $\mu = E(X)$ and $\sigma = \sqrt{Var(X)}$. Here we assume that $\mu \neq 0$.

- CV measures variation per unit of mean
- CV is very sensitive to small changes in μ when μ is near 0.

(iii) Measures of Skewness

Skewness of a probability distribution is a measure of its asymmetry (lack of symmetry).

Recall that:

Distribution of X is symmetric about μ

$\iff X-\mu \overset{d}{=} \mu - X$

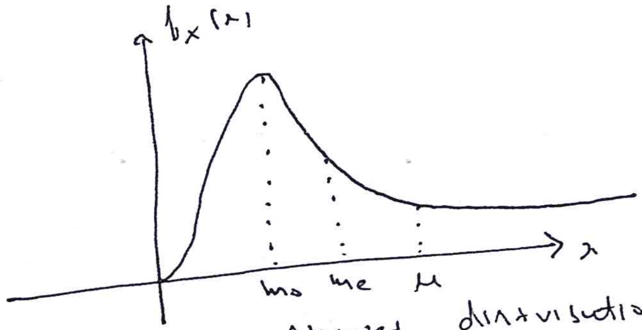
$\iff f_X(\mu+x) = f_X(\mu - x) \quad \forall x \in \mathbb{R}$

and in that case

- $\mu = E(X) =$ me (median);
- the shape of the p.d.f./p.m.f. on the left of μ is the mirror image of that on the right side of μ

Positively skewed distributions:

- Have more probability mass to the right side of p.d.f. / p.m.f.
- Have longer tails on the right side of p.d.f.



Positively skewed distribution

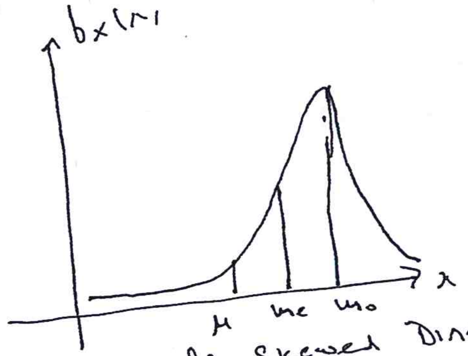
For unimodal positively skewed distribution, normally

Mode < Median < Mean

since the positive mass to large values of x pulls up the value of mean μ

Negatively skewed distributions

- Have more probability mass to the left side of the p.d.f. / p.m.f.;
- Have longer tails on the left side of p.d.f



Negatively Skewed Distribution

For unimodal negatively skewed distributions, normally

Mean < Median < Mode.

Let $\mu = E(X)$, $\sigma = \sqrt{Var(X)}$ and

$$Z = \frac{X-\mu}{\sigma} \quad : \text{ standardized variable (independent of units)}$$

Define

$$\beta_1 = E(Z^3) = \frac{E((X-\mu)^3)}{\sigma^3} = \frac{\mu_3}{\mu_2^{3/2}}, \quad \text{where } \mu_r = E((X-\mu)^r), \; r = 1, 2, \ldots$$

$\hookrightarrow$ Coefficient of skewness

- For symmetric distributions $\beta_1 = 0$. Converse may not be true.
- For positively skewed distributions, normally β_1 is large positive quantity.
- For negatively skewed distributions, normally β_1 is a small negative quantity.

A measure of skewness can also be based on quartiles.

Let

q_1: first quartile

m_e: Median (or second quartile q_2)

q_3: third quartile

μ: mean

- For symmetric distributions: $q_3 - m = m - q_1$ $\left(m = \frac{q_1+q_3}{2}\right)$
- For positively skewed distributions: $q_3 - m > m - q_1$
- For negatively skewed distributions: $q_3 - m < m - q_1$

Thus a measure of skewness can be based on $(q_3 - m) - (m - q_1) = q_3 - 2m + q_1$

Define

$$\beta_2 = \frac{(q_3 - m) - (m - q_1)}{q_3 - q_1} = \frac{q_3 - 2m + q_1}{q_3 - q_1} \quad \left(\text{independent of units}\right)$$

$\rightarrow$ Yule coefficient of skewness.

Clearly for positively/negatively skewed distribution $\beta_2 > 0 / \beta_2 < 0$ and for symmetric distributions $\beta_2 = 0$.

(IV) Measures of Kurtosis

For $\mu \in \mathbb{R}$ and $\sigma > 0$, let $Y_{\mu,\sigma}$ be a r.v. having p.d.f

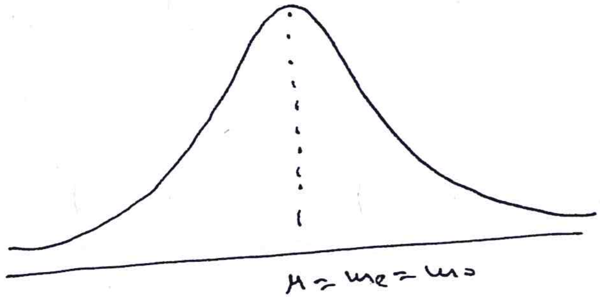
$$f_{Y_{\mu,\sigma}}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

$\rightarrow$ Normal distribution $(Y_{\mu,\sigma} \sim N(\mu, \sigma^2))$

It can be shown that

- $E(Y_{\mu,\sigma}) = \mu$;
- $Var(Y_{\mu,\sigma}) = \sigma^2$;
- $Y_{\mu,\sigma} - \mu \stackrel{d}{=} \mu - Y_{\mu,\sigma}$ and hence $\beta_1 = 0$
- $E((Y_{\mu,\sigma} - \mu)^4) = 3\sigma^4$.

- $f_{Y_{\mu,\sigma}}(\cdot)$ is unimodal & symmetric.



$\mu = me = mo$

Kurtosis of the probability distribution of X is a measure of peakedness and thickness of tails of p.m.f./p.d.f. of X relative to that of normal distribution

A distribution is said to have higher (lower) kurtosis than the normal distribution if its p.m.f./p.d.f., in comparison with p.d.f. of a normal distribution, has a sharper (rounded) peak and longer, fatter (shorter, thinner) tails.

Define $Z = (X-\mu)/\sigma$ (independent of units)

$$\gamma_1 = E(Z^4) = \frac{E((X-\mu)^4)}{\sigma^4} = \frac{\mu_4}{\mu_2^2}$$

$\hookrightarrow$ Kurtosis of the probability distribution of X

γ_1 is used as a measure of kurtosis for unimodal distributions

For $N(\mu, \sigma^2)$ distribution, $\gamma_1 = 3$.

The quantity

$$\gamma_2 = \gamma_1 - 3$$

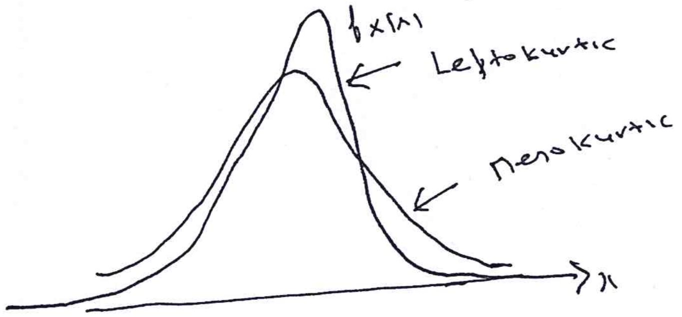
is called the excess kurtosis of the distribution of X.

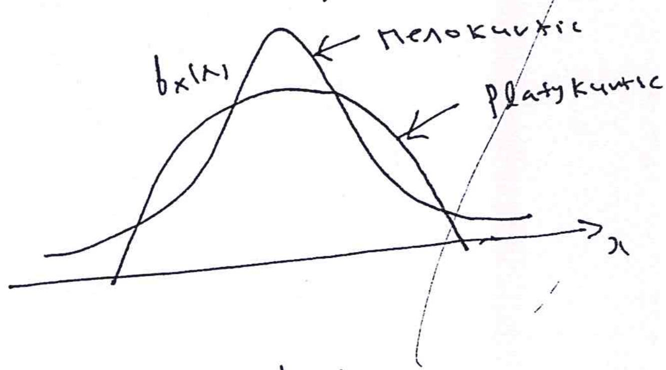
Obviously for normal distributions $\gamma_2 = 0$.

Mesokurtic distributions: Distributions with $\gamma_2 = 0$

Leptokurtic distributions: Distributions with $\gamma_2 > 0$
(has sharper peak and longer, fatter tails)

Platykurtic distributions: Distributions with $\gamma_2 < 0$
(has rounded peak and shorter, thinner tails)





Example 5.1.1 For $\alpha \in [0,1)$, let X_α have the p.d.f.

$$f_\alpha(x) = \begin{cases} \alpha e^{x}, & x < 0 \\ (1-\alpha) e^{-x}, & x \geq 0 \end{cases}$$

Recall that for $r \in \{1, 2, \dots\}$

$$I_r = \int_0^\infty e^{-x} x^{r-1} \, dx = \lfloor\underline{r-1} \quad \text{(using integration by parts)}$$

Thus, for $r \in \{1, 2, \dots\}$

$$\mu_r'(\alpha) = E(X_\alpha^r) = \int_{-\infty}^{0} \alpha x^r e^{x} \, dx + \int_0^\infty (1-\alpha) x^r e^{-x} \, dx$$

$$= [(-1)^r \alpha + 1 - \alpha] \int_0^\infty x^r e^{-x} \, dx$$

$$= \begin{cases} (1-2\alpha) \lfloor\underline{r}, & r \in \{1, 3, 5, \dots\} \\ \lfloor\underline{r}, & r \in \{2, 4, 6, \dots\} \end{cases}$$

Let ξ_p be the quantile of order $p \in (0,1)$. Then
$F_\alpha(\xi_p) = p$, where F_α is the d.f. of X_α.

Clearly

$$F_\alpha(0) = \alpha \int_{-\infty}^{0} e^{x}\,dx = \alpha$$

For $0 \le \alpha < p$

$$p = F_\alpha(\xi_p)$$

$$= \int_{-\infty}^{0} \alpha e^{x}\,dx + \int_{0}^{\xi_p} (1-\alpha)\, e^{-x}\,dx$$

$$= 1-(1-\alpha)\,e^{-\xi_p}$$

and for $\alpha \ge p$

$$p = \int_{-\infty}^{\xi_p} \alpha e^{x}\,dx = \alpha e^{\xi_p}$$

Thus

$$\xi_p = \begin{cases} \ln\left(\frac{1-\alpha}{1-p}\right), & \text{if } 0 \le \alpha < p \\ -\ln\left(\frac{\alpha}{p}\right), & \text{if } p \le \alpha \le 1 \end{cases}$$

$$q_1(\alpha) = \xi_{\frac{1}{4}} = \begin{cases} \ln\left(\frac{4(1-\alpha)}{3}\right), & \text{if } 0 \le \alpha < \frac{1}{4} \\ -\ln(4\alpha), & \text{if } \frac{1}{4} \le \alpha \le 1 \end{cases}$$

$$m_e(\alpha) = \xi_{\frac{1}{2}} = \begin{cases} \ln(2(1-\alpha)), & \text{if } 0 \le \alpha < \frac{1}{2} \\ -\ln(2\alpha), & \text{if } \frac{1}{2} \le \alpha \le 1 \end{cases}$$

$$q_3(\alpha) = \xi_{\frac{3}{4}} = \begin{cases} \ln(4(1-\alpha)), & \text{if } 0 \le \alpha < 3/4 \\ -\ln\left(\frac{4\alpha}{3}\right), & \text{if } \frac{3}{4} \le \alpha \le 1 \end{cases}$$

$$\mu_1'(\alpha) = E(X_\alpha) = 1-2\alpha$$

$$\text{Mode} = m_o(\alpha) = \sup\{f_\alpha(x) : -\infty < x < \infty\} = \max\{\alpha, 1-\alpha\}$$

$$\mu_2'(\alpha) = E(X_\alpha^2) = 2$$

$$\sigma(\alpha) = \sqrt{Var(X_\alpha)} = \sqrt{1+4\alpha-\alpha^2}$$

$\boxed{14/5}$

Note that, for $0 \le \alpha < \frac{1}{2}$, $me(\alpha) = \ln(2(1-\alpha)) \ge 0$ and for $\alpha > \frac{1}{2}$, $me(\alpha) = -\ln(2\alpha) < 0$. Thus, for $0 \le \alpha < \frac{1}{2}$ (so that $me(\alpha) \ge 0$)

$$
\begin{aligned}
MD(me(\alpha)) &= E(|X - me(\alpha)|) \\
&= \alpha \int_{-\infty}^{0} (me(\alpha) - x)\, e^{x}\, dx + (1-\alpha) \int_{0}^{me(\alpha)} (me(\alpha) - x)\, e^{-x}\, dx \\
&\quad + (1-\alpha) \int_{me(\alpha)}^{\infty} (x - me(\alpha))\, e^{-x}\, dx \\
&= me(\alpha) + 2\alpha \\
&= \ln(2(1-\alpha)) + 2\alpha
\end{aligned}
$$

Similarly, for $\frac{1}{2} \le \alpha \le 1$ (so that $me(\alpha) \le 0$)

$$
\begin{aligned}
MD(me(\alpha)) &= E(|X - me(\alpha)|) \\
&= \alpha \int_{-\infty}^{me(\alpha)} (me(\alpha) - x)\, e^{x}\, dx + \alpha \int_{me(\alpha)}^{0} (x - me(\alpha))\, e^{x}\, dx \\
&\quad + (1-\alpha) \int_{0}^{\infty} (x - me(\alpha))\, e^{-x}\, dx \\
&= 2(1-\alpha) - me(\alpha) \\
&= \ln(2\alpha) + 2(1-\alpha).
\end{aligned}
$$

Thus

$$
MD(me(\alpha)) = \begin{cases} \ln(2(1-\alpha)) + 2\alpha, & \text{if } 0 \le \alpha < \frac{1}{2} \\ \ln(2\alpha) + 2(1-\alpha), & \text{if } \frac{1}{2} \le \alpha \le 1 \end{cases}
$$

$$
IQR = IQR(\alpha) = q_3(\alpha) - q_1(\alpha) = \begin{cases} \ln 3, & \text{if } 0 \le \alpha < \frac{1}{4} \text{ or } \frac{3}{4} \le \alpha \le 1 \\ \ln(16\alpha(1-\alpha)), & \text{if } \frac{1}{4} \le \alpha < \frac{3}{4} \end{cases}
$$

$$QD \equiv QD(\alpha) = \frac{q_3(\alpha) - q_1(\alpha)}{2}$$

$$= \begin{cases} \ln\sqrt{3}, & \text{if } 0 \le \alpha < \frac{1}{4} \\ \ln\left(4\sqrt{\alpha(1-\alpha)}\right), & \text{if } \frac{1}{4} \le \alpha < \frac{3}{4} \\ \ln\sqrt{3}, & \text{if } \frac{3}{4} \le \alpha \le 1 \end{cases}$$

$$CQD \equiv CQD(\alpha) = \frac{q_3(\alpha) - q_1(\alpha)}{q_3(\alpha) + q_1(\alpha)}$$

$$= \begin{cases} \dfrac{\ln 3}{\ln\left(\frac{16(1-\alpha)^2}{3}\right)}, & \text{if } 0 \le \alpha < \frac{1}{4} \\ \dfrac{\ln(16\alpha(1-\alpha))}{\ln\left(\frac{1-\alpha}{\alpha}\right)}, & \text{if } \frac{1}{4} \le \alpha \le \frac{3}{4} \\ -\dfrac{\ln 3}{\ln\left(\frac{16\alpha^2}{3}\right)}, & \text{if } \frac{3}{4} \le \alpha \le 1 \end{cases}$$

For $\alpha \ne \frac{1}{2}$

$$CV \equiv CV(\alpha) = \frac{\sigma(\alpha)}{\mu_1'(\alpha)} = \frac{\sqrt{1+4\alpha-4\alpha^2}}{1-2\alpha}$$

$$\mu_3(\alpha) = E\left((X_\alpha - \mu_1'(\alpha))^3\right)$$

$$= \mu_3'(\alpha) - 3\mu_1'(\alpha)\mu_2'(\alpha) + 2(\mu_1'(\alpha))^3$$

$$= 2(1-2\alpha)^3$$

$$\beta_1 \equiv \beta_1(\alpha) = \frac{\mu_3(\alpha)}{\sigma(\alpha)} = \frac{2(1-2\alpha)^3}{\sqrt{1+4\alpha-4\alpha^2}}$$

$$\beta_2 \equiv \beta_2(\alpha) = \frac{q_3(\alpha) - 2m(\alpha) + q_1(\alpha)}{q_3(\alpha) - q_1(\alpha)}$$

$$
= \begin{cases} \dfrac{\ln(\frac{4}{3})}{\ln 3}, & \text{if } 0 \le \alpha < \frac{1}{4} \\[2ex] -\dfrac{\ln(4\alpha(1-\alpha))}{\ln(16\alpha(1-\alpha))}, & \text{if } \frac{1}{4} \le \alpha < \frac{1}{2} \\[2ex] \dfrac{\ln(4\alpha(1-\alpha))}{\ln(16\alpha(1-\alpha))}, & \text{if } \frac{1}{2} \le \alpha \le \frac{3}{4} \\[2ex] \dfrac{\ln(\frac{3}{4})}{\ln 3}, & \text{if } \frac{3}{4} \le \alpha \le 1 \end{cases}
$$

Clearly, for $0 \le \alpha < \frac{1}{2}$, $\beta_i(\alpha) > 0$, $i = 1, 2$ and, for $\frac{1}{2} < \alpha \le 1$, $\beta_i(\alpha) < 0$, $i = 1, 2$. For $\alpha = \frac{1}{2}$, $\beta_i(\alpha) = 0$, $i = 1, 2$.

Thus

- for $0 \le \alpha < \frac{1}{2}$, distribution of X_α is positively skewed
- for $\frac{1}{2} < \alpha \le 1$, distribution of X_α is negatively skewed
- for $\alpha = \frac{1}{2}$, distribution of X_α is symmetric (in fact in this case $f_\alpha(x) = f_\alpha(-x)$ $\forall x \in \mathbb{R}$)

$$
\begin{aligned}
\mu_4 \equiv \mu_4(\alpha) &= E((X - \mu_1'(\alpha))^4) \\
&= \mu_4'(\alpha) - 4\mu_1'(\alpha)\mu_3'(\alpha) + 6(\mu_1'(\alpha))^2\mu_2'(\alpha) - 3(\mu_1'(\alpha))^4 \\
&= 24 - 12(1-2\alpha)^2 - 3(1-2\alpha)^4
\end{aligned}
$$

$$
\gamma_1 \equiv \gamma_1(\alpha) = \frac{\mu_4(\alpha)}{(\mu_2(\alpha))^2} = \frac{24 - 12(1-2\alpha)^2 - 3(1-2\alpha)^4}{[2 - (1-2\alpha)^2]^2}
$$

and $$\gamma_2 \equiv \gamma_2(\alpha) - 3 = \frac{12 - 6(1-2\alpha)^4}{(2 - (1-2\alpha)^2)^2}$$

Clearly, for any $\alpha \in [0, 1]$, $\gamma_2(\alpha) > 0$. It follows that for any value of $\alpha \in [0, 1]$ the distribution of X_α is leptokurtic.

5.2. Some Special Discrete Distributions

5.2.1. Bernoulli and Binomial distribution

Bernoulli Experiment: A random experiment with just two possible outcomes (say success (S) and failure (F)). Each replication of a Bernoulli experiment is called a Bernoulli trial.

Consider a sequence of n independent Bernoulli trials with probability of success (S) in each trial as $p \in (0, 1)$ (same for each trial); here $n \in \mathbb{N}$ is a fixed natural number.

Define

$$X = \# \text{ of successes in } n \text{ trials.}$$

Then $S_X = \{0, 1, 2, \ldots, n\}$ and for $k \in S_X$

$$P(X = k) = P(\underbrace{SS\ldots SFF\ldots F}_{k\,S \text{ and } (n-k)\,F}) + P(\underbrace{SFFS\ldots FFS}_{k\,S \text{ and } (n-k)\,F}) + \ldots + P(\underbrace{FF\ldots FSS\ldots S}_{k\,S \text{ and } (n-k)\,F}) \quad \left(\text{total of } \binom{n}{k} \text{ terms}\right)$$

$$= p^k (1-p)^{n-k} + p^k (1-p)^{n-k} + \ldots + p^k (1-p)^{n-k}$$

(independence of trials)

$$= \binom{n}{k} p^k (1-p)^{n-k}$$

Thus

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0, 1, 2, \ldots, n \\ 0, & \text{otherwise} \end{cases}$$

$\rightarrow$ Binomial distribution with n trials and success probability p (denoted by $\text{Bin}(n, p)$ and written as $X \sim \text{Bin}(n, p)$)

18/5

$\{ Bin(n, p) : n \in \mathbb{N}, p \in (0,1) \}$

$\rightarrow$ · family of probability distributions.
· has two parameters $n \in \mathbb{N}$ and $p \in (0,1)$

$\{ Bin(1, p) : p \in (0,1) \}$: Bernoulli distributions

$Bin(1, p)$: Bernoulli distribution with success probability $p \in (0,1)$.

M.g.f. Suppose that $X \sim Bin(n, p)$ $n \in \mathbb{N}$, $p \in (0,1)$.

$$M_X(t) = E(e^{tX}) = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k (1-p)^{n-k}$$

$$\boxed{M_X(t) = (1-p+pe^t)^n, \quad t \in \mathbb{R}}$$

Let $q = 1-p$, so that $M_X(t) = (q+pe^t)^n$, $t \in \mathbb{R}$.

$$M_X^{(1)}(t) = n(q+pe^t)^{n-1} pe^t$$

$$M_X^{(2)}(t) = np(q+pe^t)^{n-1} e^t + n(n-1)(q+pe^t)^{n-2}(pe^t)^2$$

$$E(X) = M_X^{(1)}(0) = np$$

$$E(X^2) = M_X^{(2)}(0) = np + n(n-1)p^2$$

$$Var(X) = E(X^2) - (E(X))^2 = np(1-p) = npq$$

Note that if $X \sim Bin(n, p)$ then

Variance < Mean

It can be seen that

$$\mu_3' = E(X^3) = np[1-3p+3np+2p^2-3np^2+n^2p^2]$$

$$\mu_4' = E(X^4) = np[1-7p+7np+12p^2-18np^2+6n^2p^2-6p^3+11np^3 - 6n^2p^3 + n^3p^3]$$

$$\mu_3 = E((X-\mu_i')^3) = np(1-p)(1-2p)$$

$$\mu_4 = E((X-\mu_i')^4) = np(1-p)\left[3p^2(2-n) + 3p(n-2) + 1\right]$$

$$\beta_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{1-2p}{\sqrt{np(1-p)}} \quad \left(\begin{array}{l}\text{Symmetric for } p = 1/2 \\ \text{positively skewed for } 0<p<\frac{1}{2} \\ \text{negatively skewed for } p > 1/2\end{array}\right)$$

$$\gamma_2 = \beta_2 - 3 = \frac{1-6pq}{npq}, \quad \text{where } \beta_2 = \frac{\mu_4}{\mu_2^2}.$$

Alt. For $r \in \{1, 2, \dots\}$, let $X_{(r)} = X(X-1)\cdots(X-r+1)$. Then

$$E(X_{(r)}) = \sum_{k=0}^{n} k(k-1)\cdots(k-r+1)\binom{n}{k}p^k(1-p)^{n-k}$$

$$= n(n-1)\cdots(n-r+1)\sum_{k=r}^{n}\binom{n-r}{k-r}p^k(1-p)^{n-k}$$

$$= n(n-1)\cdots(n-r+1)\,p^r\sum_{k=0}^{n-r}\binom{n-r}{k}p^k(1-p)^{n-r-k}$$

$$= n(n-1)\cdots(n-r+1)\,p^r(1-p+p)^{n-r}$$

$$= n(n-1)\cdots(n-r+1)\,p^r$$

Theorem 5.2.1. Let $X_1, \dots, X_k$ be independent rvs with $X_i \sim Bin(n_i, p)$, $n_i \in \mathbb{N}$, $p \in (0,1)$, $i = 1, \dots, k$. Then $Y = \sum_{i=1}^{k} X_i \sim Bin(n, p)$, where $n = \sum_{i=1}^{k} n_i$.

Proof. For $t \in \mathbb{R}$

$$M_Y(t) = E(e^{tY}) = E\left(e^{t\sum_{i=1}^{k} X_i}\right)$$

$$= E\left(\prod_{i=1}^{k} e^{tX_i}\right) = \prod_{i=1}^{k} E(e^{tX_i}) \quad (\text{independence})$$

$$= \prod_{i=1}^{k} M_{X_i}(t) = \prod_{i=1}^{k}(1-p+pe^t)^{n_i}$$

$$= (1-p+pe^t)^{\sum_{i=1}^{k} n_i}$$

$\longrightarrow$ m.g.f. of $Bin\left(\sum_{i=1}^{k} n_i, p\right)$

20/5

By uniqueness of m.g.f $Y \sim Bin(n, p)$ where $n = \sum_{i=1}^{k} n_i$.

Example 5.2.1.1 Let $X \sim Bin(n, \frac{1}{2})$. Then $X - \frac{n}{2} \stackrel{d}{=} \frac{n}{2} - X$, since $n - X \stackrel{d}{=} X$. (Exercise)

Example 5.2.1.2 A fair dice is rolled 5 times independently. Find the probability that on 3 occasions we get a six.

Solution Consider getting a six as success. Then

$X =$ # of successes in 5 trials

$\sim Bin(5, \frac{1}{6})$

Required probability $= P(X = 3)$

$= \binom{5}{3}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2.$

5.2.2. Negative Binomial Distribution

Consider a sequence of independent Bernoulli trials with probability of success in each trial as $p \in (0,1)$. Let $r \in \{1, 2, \ldots\}$ be a fixed positive integer.

$X =$ # of failures preceding the r-th success

Then $S_X = \{0, 1, 2, \ldots\}$ and for $k \in S_X$

$f_X(k) = P(X = k)$

$= P(k \text{ failures precede } r\text{-th success})$

$= P(r-1 \text{ successes in first } k+r-1 \text{ trials and success in } (k+r)\text{-th trial})$

$= P(r-1 \text{ successes in first } k+r-1 \text{ trials}) \times P(\text{success in } (k+r)\text{-th trial})$

(independence of trials)

$$= \left\{ \binom{k+r-1}{r-1} p^{r-1} (1-p)^k \right\} \times p$$

$$= \binom{k+r-1}{r-1} p^r (1-p)^k.$$

Thus

$$f_X(x) = P(X=x) = \begin{cases} \binom{x+r-1}{r-1} p^r (1-p)^x, & x = 0, 1, 2, \dots \\ 0, & \text{otherwise} \end{cases}$$

$\longrightarrow$ Negative binomial distribution with r failures and success probability $p \in (0,1)$ (denoted by $NB(r, p)$, and written as $X \sim NB(r, p)$) (has two parameters $r \in \mathbb{N}$ and $p \in (0,1)$)

$\{NB(r, p) : r \in \mathbb{N}, p \in (0,1)\}$ is a family of probability distributions

Remark S.2.2.1: For $t \in (-1, 1)$

$$\sum_{k=0}^{\infty} \binom{k+r-1}{r-1} t^k = 1 + rt + \frac{(r+1)r}{\lfloor 2} t^2 + \frac{(r+2)(r+1)r}{\lfloor 3} t^3 + \dots$$

$$= (1-t)^{-r}.$$

The m.g.f. of $X \sim NB(r, p)$ is

$$M_X(t) = E(e^{tX}) = \sum_{k=0}^{\infty} e^{kt} \binom{k+r-1}{r-1} p^r (1-p)^k$$

$$= p^r \sum_{k=0}^{\infty} \binom{k+r-1}{r-1} ((1-p)e^t)^k \qquad t < -\ln(1-p)$$

$$= p^r (1 - (1-p)e^t)^{-r},$$

$$\Rightarrow \boxed{\Pi_X(t) = \left(\frac{p}{1-(1-p)e^t}\right)^r, \quad t < -\ln(1-p)}$$

$$\psi_X(t) = \ln \Pi_X(t) = r\ln p - r\ln(1-qe^t), \quad t < -\ln(1-p)$$

$$\psi_X^{(1)}(t) = \frac{rqe^t}{1-qe^t} = r\left[\frac{1}{1-qe^t} - 1\right], \quad t < -\ln(1-p)$$

$$\psi_X^{(2)}(t) = \frac{rqe^t}{(1-qe^t)^2}, \quad t \in \mathbb{R}$$

$$E(X) = \psi_X^{(1)}(0) = \frac{rq}{p}$$

$$\operatorname{var}(X) = \psi_X^{(2)}(0) = \frac{rq}{p^2}$$

Variance > mean

Alt For $m \in \{1, 2, \dots\}$ let $X_{(m)} = X(X-1)\cdots(X-m+1)$. Then

$$E(X_{(m)}) = \sum_{k=0}^{\infty} k(k-1)\cdots(k-m+1)\binom{k+r-1}{r-1} p^r (1-p)^k$$

$$= p^r \sum_{k=m}^{\infty} k(k-1)\cdots(k-m+1)\binom{k+r-1}{r-1}(1-p)^k$$

$$= r(r+1)\cdots(r+m-1)\,p^r \sum_{k=m}^{\infty} \frac{\lfloor k+r-1}{\lfloor k-m \ \lfloor r+m-1} (1-p)^k$$

$$= r(r+1)\cdots(r+m-1)\,p^r \sum_{k=0}^{\infty} \frac{\lfloor k+m+r-1}{\lfloor k \ \lfloor r+m-1} (1-p)^{k+m}$$

$$= r(r+1)\cdots(r+m-1)\,p^r q^m \sum_{k=0}^{\infty} \binom{k+m+r-1}{m+r-1} q^k$$

$$= r(r+1)\cdots(r+m-1)\,p^r q^m (1-q)^{-(m+r)}$$

$$= r(r+1)\cdots$$

$$= r(r+1)\cdots(r+m-1)\left(\frac{q}{p}\right)^m$$

$$\mu_1' = E(X) = \frac{rq}{p};$$

$$\mu_2' = E(X^2) = \frac{rq(1+rq)}{p^2};$$

It can be seen that

$$\mu_3' = E(X^3) = \frac{q[rp^2 + 3pqr + q^2 r(r+1)]}{p^3};$$

$$\mu_4' = E(X^4) = \frac{q[rp^3 + 7p^2 qr + 6pq^2 r(r+1) + q^3 r(r+1)(r+2)]}{p^4};$$

$$\mu_2 = E((X-\mu_1')^2) = r(1-p);$$

$$\mu_3 = E((X-\mu_1')^3) = \frac{r(p-1)(p-2)}{p^3};$$

$$\mu_4 = E((X-\mu_1')^4) = \frac{r(1-p)(6-6p+p^2+3r-3pr)}{p^4};$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{2-p}{\sqrt{rq}} > 0 \quad (\text{Positively skewed})$$

$$\delta_2 = \delta_1 - 3 = \frac{p^2-2p+6}{rq}, \quad \text{where } \delta_1 = \frac{\mu_4}{\mu_2^2}.$$

$NB(1,p)$ distribution is called a geometric distribution (denoted by $Ge(p)$), $0<p<1$. The p.m.f of $Y \sim Ge(p)$ is given by

$$f_Y(y) = P(Y=y) = \begin{cases} pq^y, & y = 0,1,2,\ldots \\ 0, & \text{otherwise} \end{cases}$$

$$P(Y \geq m) = p\sum_{y=m}^{\infty} q^y = q^m$$

$$\Rightarrow P(Y \geq m+n \mid Y \geq m) = \frac{P(Y \geq m+n, Y \geq m)}{P(Y \geq m)}$$

$$= \frac{P(Y \geq m+n)}{P(Y \geq m)}$$

$$= \frac{q^{m+n}}{q^m} = q^n$$

$$= P(Y \geq n) \quad \forall\, m, n \in \{0, 1, 2, \ldots\} \quad \ldots\ldots (A)$$

$$\Leftrightarrow P(Y \geq m+n) = P(Y \geq m)\, P(Y \geq n), \quad \forall\, m, n \in \{0, 1, 2, \ldots\} \quad \ldots (B)$$

Remark 5.2.2 The property (A) possessed by $Ge(p)$ distribution has an interesting interpretation. Suppose that a device can absorb 0, 1, 2, ... shocks before failing. Let T denote the random variable representing the number of shocks that device can absorb before failing.

$P(T \geq m+n \mid T \geq m)$: Conditional probability that a system that has absorbed m shocks will absorb at least n additional shocks before failing

$P(T \geq n)$: a new device can survive at least n shocks before failing

Thus if distribution of T has property (A) then the age of the device has no effect on the residual (remaining) life of the device (implying that an used device is as good as new as a new device). The property (A) (or equivalently (B)) is famously known as Lack of Memory Property

Theorem 5.2.2.1 Let T be a discrete type rv with range $S_T = \{0, 1, 2, \ldots\}$. Then T has the lack of memory property if and only if $T \sim Ge(p)$, for some $p \in (0, 1)$.

Proof Obviously

$T \sim Ge(p)$ for some $p \in (0,1) \Rightarrow T$ has LOM property.

Conversely suppose that T has LOM property. Then

$$P(T \geq j+k) = P(T \geq j)\, P(T \geq k), \quad \forall\, j, k \in \{0, 1, 2, \ldots\}$$

Let $P(T=0) = p$. Then $p \in (0, 1)$ and, for $j \in \{0, 1, 2, \ldots\}$

$$\begin{aligned} P(T \geq j+1) &= P(T \geq j)\, P(T \geq 1) \\ &= P(T \geq j)\,(1-p) \\ &= P(T \geq j-1)\,(1-p)^2 \\ &\;\vdots \\ &= P(T \geq 0)\,(1-p)^{j+1} \\ &= (1-p)^{j+1} \end{aligned}$$

$$\Rightarrow \quad P(T=k) = P(T \geq k) - P(T \geq k+1)$$
$$= p(1-p)^k, \quad k = \{0, 1, 2, \ldots\}.$$

$$\Rightarrow \quad T \sim Ge(p).$$

Example 5.2.2.1 A person repeatedly rolls a fair dice independently until an upper face with two or three dots is observed twice. Find the probability that the person would require eight rolls to achieve this.

Solution Consider getting 2 or 3 dots as success. Let

$Z =$ # of trials required to get 2 successes.

Then probability of success in each trial is $1/3$ and

Required probability $= P(Z=8)$

$$= \left\{\binom{7}{1}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^6\right\} \times \frac{1}{3}$$

$$= \frac{448}{6561}.$$

5.2.3. The Hypergeometric Distribution

Consider a population comprising of $N (\geq 2)$ units out of which $a \in \{1, 2, \ldots, N-1\}$ are labeled as S (success) and $N-a$ are labeled as F (failure). A sample of size n is drawn from this population drawing one unit at a time. Let

$$X = \# \text{ of successes in drawn sample}$$

Case I. Draws are independent and sampling is with replacement (i.e. after each draw the drawn unit is replaced back into the population).

In this case we have sequence of n independent Bernoulli trials with probability of success in each trial as $p = \frac{a}{N}$. Thus

$$X \sim \text{Bin}\left(n, \frac{a}{N}\right).$$

Case II. Sampling is without replacement (i.e., drawn units are not replaced back into the population)

Here

$$P(\text{obtaining S in first draw}) = \frac{a}{N}$$

$$P(\text{obtaining S in second draw}) = \frac{a}{N} \cdot \frac{a-1}{N-1} + \frac{N-a}{N} \cdot \frac{a}{N-1}$$

$$= \frac{a}{N}$$

In general

$$P(\text{obtaining S in } i\text{-th trial}) = \frac{a}{N}, \quad i = 1, \ldots, n$$

(Exercise)

$$P(\text{obtaining S in first and second trial})$$

$$= \frac{a}{N} \cdot \frac{a-1}{N}$$

$$\neq \frac{a}{N} \cdot \frac{a}{N} = P(\text{obtaining S in first trial}) \times P(\text{obtaining S in second trial})$$

$\Rightarrow$ Draws are not independent

Thus we can not conclude that $X \sim Bin(n, \frac{a}{N})$.

$$f_X(x) = P(X=x) = \begin{cases} \dfrac{\binom{a}{x}\binom{N-a}{n-x}}{\binom{N}{n}} & , x = \max\{0, n-N+a\}, \ldots, \min\{n, a\} \\ 0 & , \text{Otherwise} \end{cases}$$

$\longrightarrow$ Hypergeometric distribution ($Hyp(a, n, N)$); has three parameters $N \in \{2, 3, \ldots\}$, $a, n \in \{1, 2, \ldots, N-1\}$

(family of probability distributions)

For $r \in \mathbb{N}$, let $X_{(r)} = X(X-1)\cdots(X-r+1)$. Then

$$E(X_{(r)}) = \frac{1}{\binom{N}{n}} \sum_{k=\max\{0, n-N+a\}}^{\min\{n, a\}} k(k-1)\cdots(k-r+1)\binom{a}{k}\binom{N-a}{n-k}.$$

Clearly for $r > \min\{n, a\}$ $E(X_{(r)}) = 0$. For $1 \le r \le \min\{n, a\}$

$$E(X_{(r)}) = \frac{1}{\binom{N}{n}} \sum_{k=\max\{r, n-N+a\}}^{\min\{n, a\}} k(k-1)\cdots(k-r+1)\binom{a}{k}\binom{N-a}{n-k}$$

$$= \frac{a_{(r)}}{\binom{N}{n}} \sum_{k=\max\{r, n-N+a\}}^{\min\{n, a\}} \binom{a-r}{k-r}\binom{N-a}{n-k}$$

$$= \frac{a_{(r)}}{\binom{N}{n}} \sum_{k=\max\{0, n-N+a-r\}}^{\min\{n-r, a-r\}} \binom{a-r}{k}\binom{N-a}{n-r-k}$$

$$= \frac{a_{(r)}}{\binom{N}{n}} \sum_{k=\max\{0, (n-r)-(N-r)+a-r\}}^{\min\{n-r, a-r\}} \binom{a-r}{k}\binom{(N-r)-(a-r)}{(n-r)-k}$$

$$= \frac{\binom{N-r}{n-r}}{\binom{N}{n}} a_{(r)},$$

Since
$$\sum_{k=\max\{0,\, m-M+b\}}^{\min\{m,b\}} \binom{b}{k}\binom{M-b}{m-k} = \binom{M}{m}.$$

Thus, for $r \in \mathbb{N}$

$$E(X_{(r)}) = \begin{cases} \dfrac{\binom{N-r}{n-r}}{\binom{N}{n}} a_{(r)}, & \text{if } r \le \min\{n,a\} \\ 0, & \text{if } r > \min\{n,a\} \end{cases}$$

In particular

$E(X) = E(X_{(1)}) = n\frac{a}{N} = np$ (say), where $p = \frac{a}{N}$.

$E(X(X-1)) = E(X_{(2)}) = \frac{n(n-1)}{N(N-1)} a(a-1)$

$$\begin{aligned} Var(X) &= E(X^2) - (E(X))^2 \\ &= E(X(X-1)) + E(X) - (E(X))^2 \\ &= n\left(\frac{a}{N}\right)\left(1-\frac{a}{N}\right)\frac{N-n}{N-1} \\ &= np(1-p)\left(1-\frac{n-1}{N-1}\right) \quad \ldots\ldots (*) \end{aligned}$$

Remark 5.23.1 In case of sampling with replacement we have $X \sim Bin(n,p)$, $E(X) = np$ and $Var(X) = np(1-p)$, where $p = \frac{a}{N}$. The factor $\left(1-\frac{n-1}{N-1}\right)$, which on multiplying to variance of $Bin(n,p)$ distribution yields the variance of $Hyp(a,n,N)$ distribution (see (*)) is called the finite population correction (f.p.c.). Clearly if the sample size n is significantly smaller than the population size N ($n \ll N$) then f.p.c will be close to 1 and variances of $Bin(n,p)$ and $Hyp(a,n,N)$ distributions will be very close. In fact when $n \ll N$ and $n \ll a \equiv a_N$ (say) are such that as $N \to \infty$ $a_N \to \infty$ and $\frac{a_N}{N}$ is a fixed quantity (i.e. $\frac{a_N}{N} \to p \in (0,1)$, where $p \in (0,1)$ is a fixed quantity) then $Bin(n, \frac{a}{N})$ distribution provides an approximation to $Hyp(a,n,N)$ distribution. Regarding choice of sample size n for using this approximation a guideline, based on various empirical studies, is that the sample size n should not exceed 10% of the population size N.

(Binomial Approximation to Hypergeometric Distribution)

Theorem 5.2.31 Let $X_{a_N,n,N} \sim Hyp(a_N, n, N)$, where a_N depends on N and $\lim_{N\to\infty} \frac{a_N}{N} = p \in (0,1)$. Let $f_{a_N,n,N}(\cdot)$ denote the p.m.f. of $X_{a_N,n,N}$. Then

$$\lim_{N\to\infty} f_{a_N,n,N}(k) = \lim_{N\to\infty} P(X_{a_N,n,N} = k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k}, & \text{if } k \in \{0,1,\dots,n\} \\ 0, & \text{otherwise.} \end{cases}$$

i.e., for large N and large a_N, so that $p = \frac{a_N}{N} \in (0,1)$ is a fixed quantity, $Hyp(a_N, n, N)$ probabilities can be approximated by $Bin(n, \frac{a}{N})$ probabilities.

Proof. : $S_X = \{m \in \mathbb{N} : \max\{0, n-N+a_N\} \le m \le \min\{n, a_N\}\}$, and $a_N = N\frac{a_N}{N} \to \infty$, as $N \to \infty$.

$n - N + a_N = N\left(\frac{n}{N} - 1 + \frac{a_N}{N}\right) \to \infty$

Also for $k \in S_X$,

$$f_X(k) = \binom{n}{k} \left\{ \prod_{j=0}^{k-1} \left(\frac{a_N - j}{N - j} \right) \right\} \left\{ \prod_{j=0}^{n-k-1} \left(\frac{N - a_N - j}{N - k - j} \right) \right\}$$

$$\xrightarrow{N\to\infty} \binom{n}{k} \left\{ \prod_{j=0}^{k-1} (p) \right\} \left\{ \prod_{j=0}^{n-k-1} (1-p) \right\}$$

$$= \binom{n}{k} p^k (1-p)^{n-k}.$$

$$\Rightarrow \lim_{N\to\infty} f_{a_N,n,N}(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k}, & \text{if } k \in \{0,1,\dots,n\} \\ 0, & \text{otherwise.} \end{cases}$$

The m.g.f. of $X \sim Hyp(a, n, N)$, although exists (since S_X is finite), can not be expressed in closed form.

5.2.4. <u>The Poisson Distribution</u>

Some event E (say accidents in a crossing) is occurring randomly over a period of time.

X = # of times event E has occurred in an unit interval (say $(0, 1]$)

To model probability distribution of X, partition the unit interval into a large number (say n, where $n \to \infty$) of infinitesimal subintervals $(\frac{i-1}{n}, \frac{i}{n}]$, $i = 1, \ldots, n$ of length $\frac{1}{n}$ each. In many situations, it may be relevant to assume that

(i) for each infinitesimal interval $(\frac{k-1}{n}, \frac{k}{n}]$, $k = 1, 2, \ldots, n$, the probability that E will occur in this interval is p_n and that it will not occur in this interval is $1 - p_n$; here $p_n \to 0$ as $n \to \infty$ and $np_n \to \lambda \in (0, \infty)$, as $n \to \infty$

(ii) Chance of two or more occurrences of E in any infinitesimal subinterval $(\frac{k-1}{n}, \frac{k}{n}]$, $k = 1, \ldots, n$, is so small that it can be neglected;

(iii) occurrences of E in two disjoint infinitesimal intervals are independent.

$X \equiv X_n$ = # of times event E occurs in $(0, 1]$

$\sim \text{Bin}(n, p_n)$.

The p.m.f. of X_n is

$$f_n(k) = \binom{n}{k} p_n^k (1-p_n)^{n-k} \, I_{\{0, 1, \ldots, n\}}(k)$$

$$= \frac{1}{\lfloor k} \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) (np_n)^k \left(1 - \frac{np_n}{n}\right)^{-k} I_{\{0, 1, \ldots, n\}}(k)$$

$$\to \frac{e^{-\lambda} \lambda^k}{\lfloor k} I_{\{0, 1, 2, \ldots\}}(k)$$

$$= \begin{cases} \dfrac{e^{-\lambda}\lambda^k}{\underline{|k}}, & \text{if } k \in \{0, 1, 2, \dots\} \\ 0, & \text{otherwise} \end{cases}$$

$\longrightarrow$ Poisson distribution ($Po(\lambda)$, $\lambda > 0$) (family of probability distributions)

Note that $\sum_{k=0}^{\infty} \frac{\lambda^k}{\underline{|k}} = e^{\lambda}$.

A rv X is said to have a Poisson distribution with parameter $\lambda > 0$ (written as $X \sim Po(\lambda)$) if its p.m.f. is given by

$$f_X(k) = P(X = k) = \begin{cases} \dfrac{e^{-\lambda}\lambda^k}{\underline{|k}}, & \text{if } k = 0, 1, 2, \dots \\ 0, & \text{otherwise} \end{cases}$$

Theorem 5.2.4.1 (Poisson Approximation to Binomial Distribution)

Let $X_n \sim Bin(n, p_n)$, $n = 1, 2, \dots$, where $p_n \in (0, 1)$, $n = 1, 2, \dots$ and $\lim_{n\to\infty} (np_n) = \lambda$, for some $\lambda > 0$. Then

$$\lim_{n\to\infty} f_{X_n}(k) = \lim_{n\to\infty} P(X_n = k) = \begin{cases} \dfrac{e^{-\lambda}\lambda^k}{\underline{|k}}, & k \in \{0, 1, 2, \dots\} \\ 0, & \text{otherwise} \end{cases}.$$

Proof: As above.

Remark 5.2.4.1 If n is large and p is small ($p_n \to 0$ as $n \to \infty$) so that np is a fixed quantity in $(0, \infty)$. ($np_n \longrightarrow \lambda > 0$) then Poisson distribution provides a good approximation to Binomial distribution.

Example 5.2.4.1 Consider a person who plays 2500 games (independently). If the probability of person winning any game is 0.002, find the probability that the person will win at least two games.

Solution: Let

X = # of wins (successes) in 2500 games played by person

Clearly $X \sim Bin(2500, 0.002)$, where $n = 2500$ and $np = 5$ $(= \lambda$, say$)$ is fixed. Therefore

$$P(X \geq 2) \overset{\text{approx.}}{\approx} P(Y \geq 2),$$

where $Y \sim Po(5)$.

Thus

$$P(X \geq 2) \approx 1 - (P(Y=0) + P(Y=1))$$

$$= 1 - (e^{-5} + 5e^{-5}) = 0.9596$$

Suppose that $X \sim Po(\lambda)$, for some $\lambda > 0$. Then for $r \in \{1, 2, \ldots\}$

$$E(X_{(r)}) = E(X(X-1)\cdots(X-r+1))$$

$$= \sum_{k=0}^{\infty} k(k-1)\cdots(k-r+1) \frac{e^{-\lambda}\lambda^k}{\lfloor k}$$

$$= e^{-\lambda} \sum_{k=r}^{\infty} \frac{\lambda^k}{\lfloor k-r}$$

$$= e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^{j+r}}{\lfloor j} = \lambda^r e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{\lfloor j}$$

$$= \lambda^r.$$

Thus

$$\mu_1 = E(X) = E(X_{(1)}) = \lambda$$

$$E(X^2) = E(X_{(2)}) + E(X) = \lambda^2 + \lambda$$

$$\Rightarrow Var(X) = E(X^2) - (E(X))^2 = \lambda \quad (= \sigma^2 = \mu_2)$$

$$\boxed{\text{Mean} = \text{variance}}$$

$$\mu_3' = E(X^3) = \lambda(\lambda^2 + 3\lambda + 1)$$

$$\mu_4' = E(X^4) = \lambda(\lambda^3 + 6\lambda^2 + 7\lambda + 1)$$

$$\mu_3 = \lambda; \quad \mu_4 = \lambda(3\lambda + 1)$$

$$\beta_1 = \frac{\mu_3}{\sigma^3} = \frac{1}{\sqrt{\lambda}}, \quad \beta_2 = \beta_1 - 3 = \frac{\lambda(3\lambda+1)}{\lambda^2} - 3$$

$$= \frac{1}{\lambda}.$$

$\boxed{33/5}$

$$M_X(t) = E(e^{tX})$$

$$= \sum_{k=0}^{\infty} e^{tk} \frac{e^{-\lambda} \lambda^k}{\lfloor k}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k}{\lfloor k}$$

$$= e^{-\lambda} e^{\lambda e^t}$$

$$\boxed{M_X(t) = e^{\lambda(e^t - 1)}, \quad t \in \mathbb{R}}$$

$$\psi_X(t) = \ln M_X(t) = \lambda(e^t - 1)$$

$$\psi_X^{(r)}(t) = \lambda e^t, \quad r = 1, 2, \ldots$$

$$\Rightarrow \quad E(X) = \psi_X^{(1)}(0) = \lambda$$

$$Var(X) = \psi_X^{(2)}(0) = \lambda$$

Theorem 5.2.4.2 Let $X_1, \ldots, X_k$ be independent rvs such that $X_i \sim Po(\lambda_i)$, for some $\lambda_i > 0$ $i = 1, \ldots, k$. Then $Y = \sum_{i=1}^{k} X_i \sim Po(\lambda)$, where $\lambda = \sum_{i=1}^{k} \lambda_i$.

Proof. For $t \in \mathbb{R}$

$$M_Y(t) = E(e^{tY}) = E(e^{t \sum_{i=1}^{k} X_i})$$

$$= E\left(\prod_{i=1}^{k} e^{tX_i}\right)$$

$$= \prod_{i=1}^{k} E(e^{tX_i}) \quad \text{(independence of } X_i\text{'s)}$$

$$= \prod_{i=1}^{k} M_{X_i}(t) = \prod_{i=1}^{k} e^{\lambda_i(e^t - 1)}$$

$$= e^{\lambda(e^t - 1)} \quad t \in \mathbb{R}$$

$$\Rightarrow \quad Y \sim Po(\lambda), \text{ where } \lambda = \sum_{i=1}^{k} \lambda_i$$

5.2.5. The Discrete Uniform Distribution

N: a given positive integer

$x_1 < x_2 < \cdots < x_N$: given real numbers

A r.v. X is said to follow a discrete uniform distribution on the set $\{x_1, \ldots, x_N\}$ (written as $X \sim U(\{x_1, \ldots, x_N\})$) if its p.m.f. is given by

$$f_X(x) = P(X=x) = \begin{cases} \frac{1}{N}, & x \in \{x_1, \ldots, x_N\} \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that $X \sim U(\{x_1, \ldots, x_N\})$. Then,

$$\mu_r' = E(X^r) = \frac{1}{N} \sum_{i=1}^{N} x_i^r$$

$$\text{Mean} = \mu_1' = \frac{1}{N} \sum_{i=1}^{N} x_i$$

$$Var(X) = \sigma^2 = E((X-\mu_1')^2) = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu_1')^2$$

m.g.f. $M_X(t) = E(e^{tX}) = \frac{1}{N} \sum_{i=1}^{N} e^{t x_i}$

Suppose that $Y \sim U(\{1, 2, \ldots, N\})$. Then

$$\mu_1' = \text{Mean} = E(Y) = \frac{1}{N} \sum_{i=1}^{N} i = \frac{N+1}{2}$$

$$\mu_2' = E(Y^2) = \frac{1}{N} \sum_{i=1}^{N} i^2 = \frac{(N+1)(2N+1)}{6}$$

$$\mu_3' = E(Y^3) = \frac{1}{N} \sum_{i=1}^{N} i^3 = \frac{N(N+1)^2}{4}$$

$$\mu_4' = E(Y^4) = \frac{1}{N} \sum_{i=1}^{N} i^4 = \frac{(N+1)(2N+1)(3N^2+3N-1)}{30}$$

$$\mu_2 = E((Y-\mu_1')^2) = \frac{N^2-1}{12}$$

$$\mu_3 = E((Y-\mu_1')^3) = 0$$

$$\mu_4 = E((Y-\mu_1')^4) = \frac{(3N^2-7)(N^2-1)}{240}$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0 \qquad \text{(Coefficient of skewness)}$$

$$\text{Kurtosis} = \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3}{5}\,\frac{3N^2-7}{N^2-1}$$

The m.g.f. of $Y \sim U(\{1,2,\cdots,N\})$ is given by

$$M_Y(t) = E(e^{tY}) = \frac{1}{N}\sum_{j=1}^{N} e^{jt}$$

$$= \begin{cases} \dfrac{e^t(e^{Nt}-1)}{e^t-1}, & \text{if } t \neq 0 \\ 1, & \text{if } t=0. \end{cases}$$

Example 5.2.5.1 A person has to open a lock whose key is lost among a set of N keys. Assume that out of these N keys only one can open the lock. To open the lock the person tries keys one by one by choosing, at each attempt, one of the keys at random from the unattempted keys. The unsuccessful keys are not considered for future attempts. Let Y denote the number of attempts the person will have to make to open the lock. Show that $Y \sim U(\{1,2,\cdots,N\})$ and hence find the mean and variance of the r.v. Y.

Solution. For $r \notin \{1,2,\cdots,N\}$, we have $P(Y=r)=0$. For $r \in \{1,2,\cdots,N\}$

$$P(Y=r) = \frac{N-1}{N}\cdot\frac{N-2}{N-1}\cdots\frac{N-(r-1)}{N-(r-2)}\cdot\frac{1}{N-(r-1)} = \frac{1}{N}$$

$\Rightarrow \quad Y \sim U(\{1,2,\cdots,N\})$

$\Rightarrow \quad E(Y) = \dfrac{N+1}{2}$ and $Var(Y) = \dfrac{N^2-1}{12}$.

5.3. Some Absolutely Continuous Distributions

5.3.1 Uniform or Rectangular Distribution

Let $-\infty < \alpha < \beta < \infty$. An absolutely continuous type r.v. X is said to have a uniform (or rectangular) distribution over the interval (α, β) (written as $X \sim U(\alpha, \beta)$) if its p.d.f. is given by

$$f(x|\alpha, \beta) = \begin{cases} \dfrac{1}{\beta-\alpha}, & \text{if } \alpha < x < \beta \\ 0 & \text{otherwise} \end{cases}$$

$\longrightarrow$ family of distributions $\{U(\alpha, \beta): -\infty < \alpha < \beta < \infty\}$ corresponding to different choices of α and β ($-\infty < \alpha < \beta < \infty$)

Suppose that $X \sim U(\alpha, \beta)$, for some $-\infty < \alpha < \beta < \infty$. Then

$$\mu_r' = E(X^r) = \int_\alpha^\beta \frac{x^r}{\alpha - \beta}\, dx$$

$$= \frac{\beta^{r+1} - \alpha^{r+1}}{(r+1)(\alpha - \beta)}$$

$$= \frac{\beta^r}{r+1}\left[1 + \frac{\alpha}{\beta} + \left(\frac{\alpha}{\beta}\right)^2 + \cdots + \left(\frac{\alpha}{\beta}\right)^r\right]$$

$$\mu_r = E((X - \mu_1')^r)$$

$$= \int_\alpha^\beta \left(x - \frac{\alpha+\beta}{2}\right)^r \frac{1}{\beta - \alpha}\, dx$$

$$= \int_{-\frac{\beta-\alpha}{2}}^{\frac{\beta-\alpha}{2}} \frac{t^r}{\beta-\alpha}\, dt = \begin{cases} 0 & \text{if } r = 1, 3, 5, \ldots \\ \dfrac{(\beta-\alpha)^r}{2^r (r+1)}, & \text{if } r = 2, 4, 6, \ldots \end{cases}$$

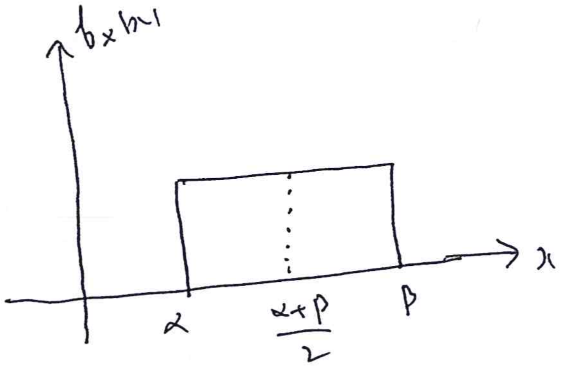
Also

$$f_X\left(x-\frac{\alpha+\beta}{2}\right)=f_X\left(\frac{\alpha+\beta}{2}-x\right)=\begin{cases}\frac{1}{\beta-\alpha}, & -\frac{\beta-\alpha}{2}<x<\frac{\beta-\alpha}{2}\\ 0, & \text{otherwise}\end{cases}$$

$$\Rightarrow \quad X-\frac{\alpha+\beta}{2} \overset{d}{=} \frac{\alpha+\beta}{2}-X$$

$\Rightarrow$ Distribution of X is symmetric about its mean

$$\mu_1'=\frac{\alpha+\beta}{2}$$



$$\text{Mean} = \mu_1' = E(X) = \frac{\alpha+\beta}{2}$$

$$\text{Var}(X) = \mu_2 = \sigma^2 = E((X-\mu_1')^2) = \frac{(\beta-\alpha)^2}{12}$$

$$\text{Coefficient of skewness} = \beta_1 = \frac{\mu_3}{\mu_2^{3/2}} = 0$$

$$\text{Kurtosis} = \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{9}{5} = 1.8$$

The d.f. of $X \sim U(\alpha,\beta)$ is given by

$$F(x|\alpha,\beta)=\begin{cases}0 & \text{if } x<\alpha\\ \frac{x-\alpha}{\beta-\alpha}, & \text{if } \alpha\le x<\beta\\ 1, & \text{if } x\ge\beta\end{cases}$$

Theorem 5.3.1.1 Let $-\infty<\alpha<\beta<\infty$ and let X be a rv of continuous type with $P(\alpha\le X\le\beta)=1$. Then

$X\sim U(\alpha,\beta) \iff P(X\in I)=P(X\in J)$, for any pairs of intervals $I, J \subseteq (\alpha,\beta)$ having the same length.

Prob. Suppose that $X \sim U(\alpha, \beta)$. Then, for $\alpha \le a < b \le \beta$

$$P(X \in (a, b)) = P(X \in [a, b)) = P(X \in (a, b]) = P(X \in [a, b])$$

$$= F(b|\alpha, \beta) - F(a|\alpha, \beta) = \frac{b-a}{\beta-\alpha}$$

$\longrightarrow$ depends only on length $b-a$ of the interval $(a,b)/[a,b)/(a,b]/[a,b]$.

Conversely suppose that

$P(X \in I) = P(X \in J)$, for all pair of intervals $I, J \subseteq (\alpha, \beta)$ having the same length.

For $0 < \lambda \le 1$, let

$$G(\lambda) = P(\alpha < X \le \alpha + (\beta-\alpha)\lambda) = F(\alpha + (\beta-\alpha)\lambda \,|\, \alpha, \beta)$$

Then, for $0 < \lambda_1, \lambda_2 \le 1$, $0 < \lambda_1 + \lambda_2 \le 1$

$$G(\lambda_1 + \lambda_2) = P(\alpha < X \le \alpha + (\beta-\alpha)(\lambda_1 + \lambda_2))$$

$$= P(\alpha < X \le \alpha + (\beta-\alpha)\lambda_1) + \underbrace{P(\alpha + (\beta-\alpha)\lambda_1 < X \le \alpha + (\beta-\alpha)(\lambda_1+\lambda_2))}_{\text{depends only on length } (\beta-\alpha)\lambda_2 \text{ of } (\alpha+(\beta-\alpha)\lambda_1,\ \alpha+(\beta-\alpha)(\lambda_1+\lambda_2))}$$

$$= G(\lambda_1) + P(\alpha < X \le \alpha + (\beta-\alpha)\lambda_2)$$

$$= G(\lambda_1) + G(\lambda_2).$$

By induction, for $0 < \lambda_i \le 1$, $i = 1, \ldots, n$, $0 < \sum_{i=1}^{n} \lambda_i \le 1$, we have

$$G(\lambda_1 + \cdots + \lambda_n) = G(\lambda_1) + \cdots + G(\lambda_n)$$

$$\Rightarrow \quad G(m\lambda) = m\,G(\lambda), \quad \forall \; 0 < \lambda \le \frac{1}{m} \quad \cdots\cdots \quad A_1$$

$$G(\lambda) = G(\underbrace{\tfrac{\lambda}{n} + \cdots + \tfrac{\lambda}{n}}_{n \text{ times}}) = n\,G(\tfrac{\lambda}{n}) \quad \cdots\cdots \quad A_2$$

For $m, n \in \{1, 2, \dots\}$, $m < n$

$$G\left(\frac{m}{n}\right) = G\Big(\underbrace{\frac{1}{n} + \cdots + \frac{1}{n}}_{m \text{ times}}\Big)$$

$$= m\, G\left(\frac{1}{n}\right) \qquad (\text{using } (A_1))$$

$$= \frac{m}{n}\, G(1) \qquad (\text{using } (A_2))$$

$$= \frac{m}{n}\, F(\beta \mid \alpha, \beta)$$

$$= \frac{m}{n}$$

$$\Rightarrow \quad G(r) = r, \qquad \forall\, r \in \mathbb{Q} \cap (0,1),$$

where $\mathbb{Q}$ denotes the set of rational numbers. Now let $\lambda \in (0,1)$. Then there exists a sequence $\{r_n\}_{n \geq 1}$ in $\mathbb{Q} \cap (0,1)$ such that $r_n \downarrow \lambda$ (rationals are dense in $(0,1)$).

Then, since G is continuous,

$$G(\lambda) = \lim_{n \to \infty} G(r_n)$$

$$= \lim_{n \to \infty} r_n$$

$$= \lambda$$

$$\Rightarrow \quad G(\lambda) = \lambda, \qquad \forall\, \lambda \in (0,1)$$

$$\Rightarrow \quad F(\alpha + (\beta - \alpha)\lambda \mid \alpha, \beta) = \lambda, \qquad \forall\, \lambda \in (0,1)$$

$$\Rightarrow \quad F(\lambda \mid \alpha, \beta) = \frac{\lambda - \alpha}{\beta - \alpha}, \qquad \forall\, \lambda \in (\alpha, \beta)$$

$$\Rightarrow \quad F(\lambda \mid \alpha, \beta) = \begin{cases} 0 & \lambda < \alpha \\ \dfrac{\lambda - \alpha}{\beta - \alpha}, & \alpha \leq \lambda < \beta \\ 1, & \lambda \geq \beta \end{cases}$$

$$\Rightarrow \quad X \sim U(\alpha, \beta).$$

m.g.f. $$M_X(t) = E(e^{tX}) = \int_{\alpha}^{\beta} \frac{e^{tx}}{\beta-\alpha}\,dx = \begin{cases} \dfrac{e^{t\beta}-e^{t\alpha}}{t(\beta-\alpha)}, & \text{if } t\neq 0 \\ 1, & \text{if } t=0 \end{cases}.$$

Theorem 5.3.1.2 Let $X \sim U(\alpha, \beta)$ $-\infty < \alpha < \beta < \infty$. Then

(i) for $a>0$ and $b \in \mathbb{R}$, $Y = aX+b \sim U(a\alpha+b,\ a\beta+b)$

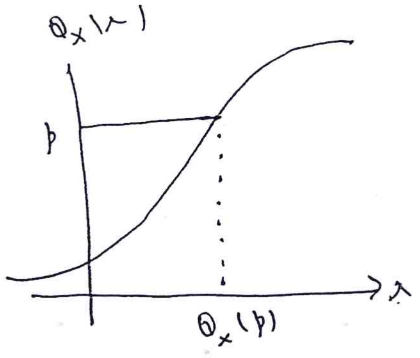
(ii) for $a<0$ and $b\in\mathbb{R}$, $Y = aX+b \sim U(a\beta+b,\ a\alpha+b)$

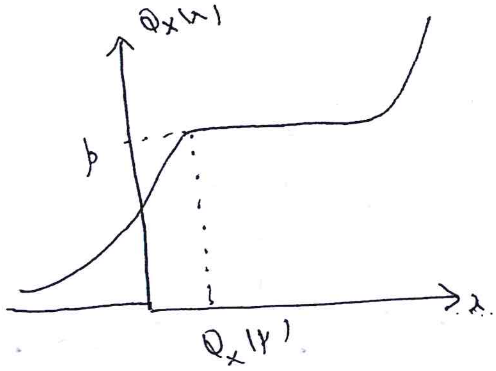
(iii) $Z = \dfrac{X-\alpha}{\beta-\alpha} \sim U(0,1)$

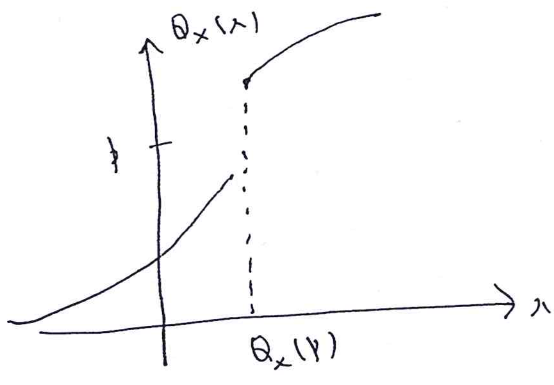
Proof. Straightforward

Recall that quantile function is defined by

$$Q_X(p) = \inf\{x \in \mathbb{R} : F_X(x) \geq p\}, \quad 0<p<1.$$







Theorem 5.3.13 Let X be a r.v. with d.f. F and quantile function $Q(\cdot)$. Then

(i) (Probability Integral Transform).

X is of continuous type $\Rightarrow$ $F(X) \sim U(0, 1)$

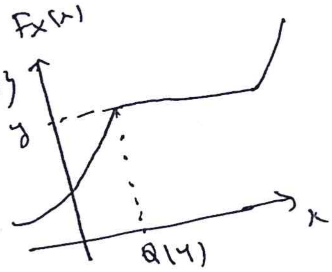
(ii) $U \sim U(0, 1) \Rightarrow Q(U) \stackrel{d}{=} X$.

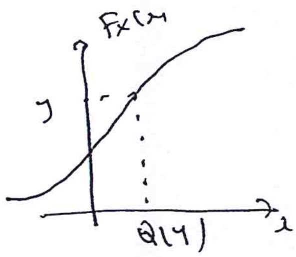
Proof (i) Let G be the d.f. of $Y = F(X)$. Then

$$G(y) = P(F(X) \le y), \quad y \in \mathbb{R}.$$

Clearly, for $y < 0$, $G(y) = 0$ and, for $y \ge 1$ $G(y) = 1$.

For $y \in [0, 1)$





$$\{s \in \mathbb{R} : F(s) \ge y\} = \{s \in \mathbb{R} : s \ge Q(y)\}$$

$$\Rightarrow P(F(X) \ge y) = P(X \ge Q(y))$$

$$\Rightarrow P(F(X) < y) = P(X < Q(y))$$

$$\Rightarrow P(F(X) < y) = P(X \le Q(y)) \quad \text{(since } X \text{ is of continuous type)}$$
$$= F(Q(y)) = y$$

Since X is of continuous type

$$P(F(X) = y) = P(x_1 \le X \le x_2) \quad \text{for some } x_1, x_2 \text{ with } F(x_1) = F(x_2)$$
$$= 0$$

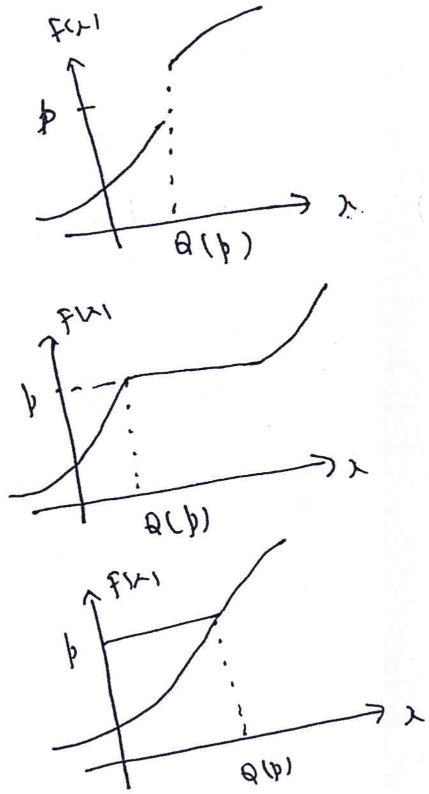
Thus

$$P(F(X) \le y) = y, \quad \forall y \in [0, 1)$$

$$\Rightarrow G(y) = \begin{cases} 0, & \text{if } y < 0 \\ y, & \text{if } 0 \le y < 1 \\ 1, & \text{if } y \ge 1 \end{cases}$$

$$\Rightarrow Y \sim U(0, 1).$$

(ii) Let $U \sim U(0,1)$ and let $Z = Q(U)$. Then the d.f. of Z is



$$H(z) = P(Z \le z)$$
$$= P(Q(U) \le z)$$
$$= P(Q(U) \le z,\ 0 < U < 1)$$

Note that, for $z \in [0,1)$

$$\{p \in \mathbb{R} : Q(p) \le z\} = \{p \in \mathbb{R} : F(z) \ge p\}$$

Thus, for $z \in [0,1)$

$$H(z) = P(F(z) \ge U,\ 0 < U < 1)$$
$$= P(U \le F(z))$$
$$= F(z)$$

$$\Rightarrow \quad Z = Q(U) \stackrel{d}{=} X.$$

Remarks 5.3.1.1 The above theorem provides a method to generate observations from any arbitrary distribution using $U(0,1)$ observations. Suppose that we require an observation X from a distribution having d.f. F and quantile function Q. To do so, the above theorem suggests that, generate an observation U from $U(0,1)$ distribution and take $X = Q(U)$.

5.3.2. Gamma and Related Distributions

Gamma Function: $\Gamma : (0,\infty) \longrightarrow (0,\infty)$

$$\Gamma\alpha = \int_0^\infty e^{-t}\, t^{\alpha-1}\, dt, \quad \alpha > 0$$

Gamma Function

$\longrightarrow$ Converges for any $\alpha > 0$

Integration by parts yields

- $\Gamma(\alpha+1) = \alpha\,\Gamma(\alpha)$, $\quad \alpha > 0$
- $\Gamma(1) = 1$
- For $n \in \mathbb{N}$, $\quad \Gamma(n) = \lfloor n-1$

$$\Gamma\left(\tfrac{1}{2}\right) = \int_0^\infty e^{-t}\, t^{-1/2}\, dt = 2\int_0^\infty e^{-x^2}\, dx$$

$$\left(\Gamma\left(\tfrac{1}{2}\right)\right)^2 = 4\int_0^\infty\int_0^\infty e^{-(x^2+y^2)}\, dx$$

$$= 4\int_0^\infty \int_0^{\pi/2} r e^{-r^2}\, d\theta\, dr \qquad (x = r\cos\theta,\ y = r\sin\theta)$$

$$= \pi$$

$$\Rightarrow \quad \Gamma\left(\tfrac{1}{2}\right) = \sqrt{\pi}$$

$$\Gamma\left(\tfrac{3}{2}\right) = \tfrac{1}{2}\,\Gamma\left(\tfrac{1}{2}\right) = \frac{\sqrt{\pi}}{2}$$

$$\Gamma\left(\tfrac{5}{2}\right) = \tfrac{3}{2}\cdot\tfrac{1}{2}\,\Gamma\left(\tfrac{1}{2}\right) = \frac{1.3}{2^2}\sqrt{\pi}$$

$$\Gamma\left(\frac{2n+1}{2}\right) = \frac{1.3.5\cdots(2n-1)}{2^n}\sqrt{\pi} = \frac{\lfloor 2n}{\lfloor n\, 4^n}\sqrt{\pi}, \quad n \in \mathbb{N}$$

Clearly

$$\boxed{\int_0^\infty e^{-\frac{x}{\theta}}\, x^{\alpha-1}\, dx = \theta^\alpha\,\Gamma(\alpha), \quad \alpha > 0,\ \theta > 0}$$

Definition 5.3.2.1 A rv X is said to have a gamma distribution with shape parameter $\alpha\ (>0)$ and scale parameter $\theta\ (>0)$ (written as $X \sim GAM(\alpha, \theta)$) if its p.d.f. is given by

$$f(x|\alpha,\theta) = \begin{cases} \dfrac{1}{\theta^\alpha\,\Gamma(\alpha)}\, e^{-\frac{x}{\theta}}\, x^{\alpha-1}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

$\longrightarrow$ family of distributions $\{GAM(\alpha,\theta),\ \alpha>0,\ \theta>0\}$

- $X \sim GAM(\alpha, \theta) \Rightarrow \frac{X}{\theta} \sim GAM(\alpha, 1)$ (θ is called a scale parameter since the distribution of $\frac{X}{\theta}$ does not depend on θ).

- The p.d.f of $Z \sim GAM(\alpha, 1)$ is

$$f(z) = \begin{cases} \frac{e^{-z} z^{\alpha-1}}{\Gamma\alpha}, & z > 0 \\ 0, & \text{otherwise} \end{cases}$$

- Let $Z \sim GAM(\alpha, 1)$. Then

$$E(Z^r) = \frac{1}{\Gamma\alpha} \int_0^\infty z^{\alpha+r-1} e^{-z} = \frac{\Gamma\alpha+r}{\Gamma\alpha}, \quad r > -\alpha, \quad \alpha > 0$$

$$= \alpha(\alpha+1)\cdots(\alpha+r-1), \quad \text{if } r \in \mathbb{N}$$

- Let $X \sim GAM(\alpha, \theta)$, $\alpha > 0$, $\theta > 0$. Then $Z = \frac{X}{\theta} \sim GAM(\alpha, 1)$

$$\Rightarrow E\left(\left(\frac{X}{\theta}\right)^r\right) = E(Z^r) = \frac{\Gamma\alpha+r}{\Gamma\alpha}$$

$$\Rightarrow E(X^r) = \frac{\Gamma\alpha+r}{\Gamma\alpha}\theta^r, \quad r > 0$$

$$= \alpha(\alpha+1)\cdots(\alpha+r-1), \quad \text{if } r \in \mathbb{N}.$$

- Mean $= \mu_1' = E(X) = \alpha\theta$
- $\mu_2' = E(X^2) = \alpha(\alpha+1)\theta^2$

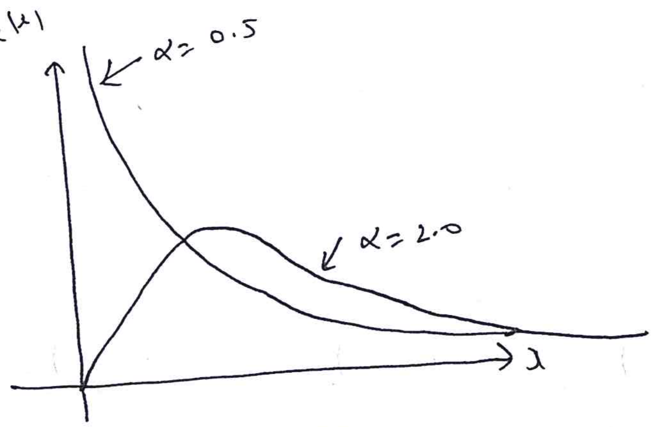
$\mu_2 = \sigma^2 = Var(X) = E(X^2) - (E(X))^2 = \alpha\theta^2$

- $\mu_3 = E((X-\mu_1')^3) = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3 = 2\alpha\theta^3$
- $\mu_4 = E((X-\mu_1')^4) = \mu_4' - 4\mu_1'\mu_3' + 6(\mu_1')^2\mu_2' - 3(\mu_1')^4$

$= 3\alpha(\alpha+2)\theta^4$

- Coefficient of skewness $= \beta_1 = \frac{\mu_3}{\mu_2^2} = \frac{2}{\sqrt{\alpha}}$
- Kurtosis $= \gamma_1 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{6}{\alpha}$

For $\alpha \le 1$, $f(x|\alpha, \theta) \downarrow$ and for $\alpha > 1$ $f(x|\alpha, \theta) \uparrow$ in $(0, (\alpha-1)\theta)$ and $\downarrow$ in $((\alpha-1)\theta, \infty)$



- m.g.f. $M_X(t) = E(e^{tX}) = E(e^{t\theta Z}) \qquad (Z = \frac{X}{\theta})$

$$= \frac{1}{\Gamma\alpha} \int_0^\infty e^{t\theta z} e^{-z} z^{\alpha-1} dz$$

$$= \frac{1}{\Gamma\alpha} \int_0^\infty e^{-(1-t\theta)z} z^{\alpha-1} dz$$

$$= (1-t\theta)^{-\alpha}, \qquad t < \frac{1}{\theta}.$$

Theorem 5.3.2.1 Let $X_1, \ldots, X_k$ be independent and let $X_i \sim GAM(\alpha_i, \theta)$, $\alpha_i > 0$, $\theta > 0$, $i = 1, \ldots, k$. Then $Y = \sum_{i=1}^{k} X_i \sim GAM(\sum_{i=1}^{k} \alpha_i, \theta)$

Proof. $M_Y(t) = \prod_{i=1}^{k} M_{X_i}(t) = \prod_{i=1}^{k} (1-t\theta)^{-\alpha_i}$

$$= (1-t\theta)^{-\sum_{i=1}^{k}\alpha_i}, \qquad t < \frac{1}{\theta}$$

$$= \text{m.g.f. of } GAM(\sum_{i=1}^{k} \alpha_i, \theta).$$

Theorem 5.3.2.2 (Relationship Between Gamma and Poisson Distribution) For $n \in \mathbb{N}$, $\theta > 0$ and $t > 0$, let $X \sim GAM(n, \theta)$ and $Y \sim Po(\frac{t}{\theta})$. Then

$$P(X > t) = P(Y \le n-1)$$

i.e. $$\frac{1}{\lfloor n-1 \, \theta^n} \int_t^\infty e^{-\frac{x}{\theta}} x^{n-1} dx = \sum_{j=0}^{n-1} \frac{e^{-t/\theta} (\frac{t}{\theta})^j}{\lfloor j} \ldots$$

Proof Use integration by parts.

Remark For $n \in \mathbb{N}$, and $\theta > 0$, let $X \sim GAM(n, \theta)$. Then

$$\sum_{j=n}^{\infty} \frac{e^{-\frac{X}{\theta}} (\frac{X}{\theta})^j}{\lfloor j} \sim U(0,1)$$

and $\sum_{j=0}^{n-1} \frac{e^{-\frac{x}{\theta}} \left(\frac{x}{\theta}\right)^j}{\lfloor j} \sim U(0,1)$ $\left(\begin{array}{l} U \sim U(0,1) \\ \Rightarrow 1-U \sim U(0,1) \end{array} \right)$

Definition 5.3.2.2 For a $\theta > 0$, a $GAM(1, \theta)$ distribution is called an exponential distribution with scale parameter θ (denoted by $Exp(\theta)$).

The p.d.f of $T \sim Exp(\theta)$ is given by

$$f_T(t) = \begin{cases} \frac{1}{\theta} e^{-t/\theta}, & t > 0 \\ 0 & \text{otherwise} \end{cases}$$

and its d.f. is given by

$$F_T(t) = \begin{cases} 0, & t < 0 \\ 1 - e^{-t/\theta}, & t \geq 0 \end{cases}$$

Mean $= E(T) = \theta$; Variance $= \mu_2 = \theta^2$

$\mu_r' = E(T^r) = \lfloor r \, \theta^r$, $r \in \mathbb{N}$

Coefficient of skewness $= \beta_1 = \frac{\mu_3}{\mu_2^{3/2}} = 2$

Kurtosis $= \gamma_1 = \frac{\mu_4}{\mu_2^2} = 9$

m.g.f. $M_T(t) = (1 - t\theta)^{-1}$, $t < \frac{1}{\theta}$

$$P(T > t) = \begin{cases} 1 & t < 0 \\ e^{-t/\theta}, & t \geq 0 \end{cases}$$

For $s > 0$, $t > 0$

$$P(T > s+t \mid T > s) = \frac{P(T > s+t)}{P(T > s)} = e^{-t/\theta} = P(T > t)$$

$\Leftrightarrow$ $P(T > s+t) = P(T > s) \, P(T > t)$, $\forall\, s, t > 0$

$\longrightarrow$ Lack of Memory Property

Let T denote the lifetime of a system. Given that the system has survived $s\ (>0)$ units of time the probability that it will survive t additional units of time is the same as the probability that a fresh unit (of age 0) will survive t units of time. In other words the system has no memory of its current age or it is not ageing with time.

Theorem 5.3.2.3 Let Y be a rv of continuous type with d.f. F such that $F(0)=0$. Then Y has LOM property (i.e. $\bar{F}(s+t)=\bar{F}(s)\bar{F}(t)$ $\forall s,t\geq 0$, where $\bar{F}=1-F$) iff $Y\sim Exp(\theta)$, for some $\theta>0$.

__Proof.__ Let $Y\sim Exp(\theta)$, $\theta>0$. Then obviously Y has LOM property. Now suppose that $F(0)=0$ and Y has LOM property. Then

$$\bar{F}(s+t)=\bar{F}(s)\bar{F}(t), \quad \forall s,t\geq 0$$

$$\Rightarrow \bar{F}(s_1+s_2+\cdots+s_m)=\bar{F}(s_1)\bar{F}(s_2)\cdots\bar{F}(s_m), \quad s_i\geq 0,\ i=1,\ldots,m$$

$$\Rightarrow \bar{F}\left(\frac{m}{n}\right)=\bar{F}(\underbrace{\tfrac{1}{n}+\cdots+\tfrac{1}{n}}_{m\text{ times}})=\left[\bar{F}\left(\tfrac{1}{n}\right)\right]^m, \quad \forall m,n\in\mathbb{N} \qquad (A_1)$$

$$\bar{F}(1)=\left[\bar{F}\left(\tfrac{1}{n}\right)\right]^n, \quad \forall n\in\mathbb{N}\ldots \qquad (A_2)$$

$$\Rightarrow \bar{F}\left(\frac{m}{n}\right)=[\bar{F}(1)]^{m/n}, \quad \forall m,n\in\mathbb{N}. \quad \ldots \quad (A_3)$$

Let $\lambda=\bar{F}(1)$ so that $0\leq\lambda\leq 1$.

$\lambda=0 \Rightarrow \bar{F}\left(\frac{1}{n}\right)=0$, $\forall n\in\mathbb{N}$ (using (A_2)) $\Rightarrow \bar{F}(0)=0$

$\Rightarrow F(0)=1$ (Contradiction, since $F(0)=0$)

$\lambda=1 \Rightarrow \bar{F}(m)=[\bar{F}(1)]^m=1$, $\forall m\in\mathbb{N}$ $\Rightarrow \lim_{m\to\infty}F(m)=0$

$\Rightarrow \lim_{m\to\infty}F(m)=0 \longrightarrow$ Contradiction

Thus $\lambda\in(0,1)$. Let $\lambda=e^{-\frac{1}{\theta}}$, $\theta>0$ $\left(\theta=-\frac{1}{\ln\lambda}\right)$

Then using (A_3)

$$\bar{F}(r)=e^{-r/\theta}, \quad \forall r\in\mathbb{Q}\cap[0,\infty)$$

Let $x\in\mathbb{Q}\cap[0,\infty)$. Then there exists a sequence $\{r_n\}_{n\geq 1}$ in $\mathbb{Q}\cap[0,\infty)$ such that $r_n\to x$. Then

$$\bar{F}(x)=\bar{F}\left(\lim_{n\to\infty}r_n\right)=\lim_{n\to\infty}\bar{F}(r_n)=\lim_{n\to\infty}e^{-r_n/\theta}=e^{-x/\theta}$$

$$\Rightarrow F(x)=\begin{cases}0, & x<0\\ 1-e^{-x/\theta}, & x\geq 0\end{cases}$$

$\Rightarrow Y\sim Exp(\theta)$.

Example 5.3.2.1 X: waiting time for occurrence of an event E

Suppose that $X \sim Exp(3)$. Then the conditional probability that the waiting time for occurrence of E is at least 5 hrs given that it has not occurred in first two hrs $= P(X>5 \mid X>2) = P(X>3) = e^{-1}$.

Chi-Squared Distribution Let $n \in \mathbb{N}$. Then $GAM(\frac{n}{2}, 2)$ distribution is called chi-squared distribution with n degrees of freedom (denoted by χ^2_n)

Let $X \sim \chi^2_n$. The pdf of X is

$$f_X(x) = \begin{cases} \frac{1}{2^{n/2}\,\Gamma\frac{n}{2}} e^{-\frac{x}{2}} x^{\frac{n}{2}-1}, & x>0 \\ 0, & \text{otherwise} \end{cases}$$

Mean $= E(X) = n$, Variance $= \mu_2 = \sigma^2 = 2n$

Coefficient of skewness $= \beta_1 = 2\sqrt{\frac{2}{n}}$

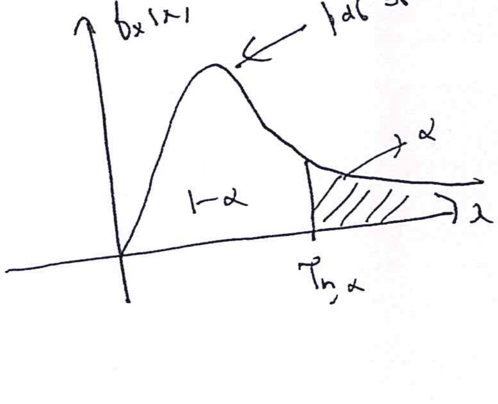
Kurtosis $= \gamma_1 = 3 + \frac{12}{n}$

m.g.f. $M_X(t) = (1-2t)^{-n/2}$, $t < \frac{1}{2}$.

Theorem 5.3.2.4 Let $X_1, \ldots, X_k$ be independent with $X_i \sim \chi^2_{n_i}$, $i=1,\ldots,k$, $n_i \in \mathbb{N}$.

Then $\sum_{i=1}^{k} X_i \sim \chi^2_n$, where $n = \sum_{i=1}^{k} n_i$

For various values of $n \in \mathbb{N}$ and $\alpha \in (0,1)$, tables for $(1-\alpha)$th quantile of χ^2_n distribution (i.e., $\tau_{n,\alpha}$ satisfying $P(\chi^2_n \le \tau_{n,\alpha}) = 1-\alpha$) are available in various textbooks



5.3.3. Beta Distribution

For $\alpha > 0$, $\beta > 0$

$$\Gamma\alpha\, \Gamma\beta = \int_0^\infty \int_0^\infty e^{-(x+t)} x^{\alpha-1} t^{\beta-1}\, dx\, dt$$

$$= \int_0^1 \int_0^\infty e^{-v} (uv)^{\alpha-1} ((1-u)v)^{\beta-1} |v| \, dv\, du$$

Making transformation:
$x = uv$
$t = (1-u)v$
$J = v$

$$= \Gamma\alpha+\beta \int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\, du$$

$$\Rightarrow \frac{\Gamma\alpha\, \Gamma\beta}{\Gamma\alpha+\beta} = \int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\, du$$

$$\| $$

$B(\alpha, \beta) \longrightarrow$ Beta function (function of (α, β), $\alpha > 0$, $\beta > 0$)

__Note:__ $B(\alpha, \beta) = B(\beta, \alpha)$, $\forall\, \alpha, \beta > 0$

5.3.3.1 Definition For given $\alpha > 0$ and $\beta > 0$, a r.v. X is said to have beta distribution with parameters (α, β) (written as $X \sim Be(\alpha, \beta)$) if the pdf of X is given by

$$f(x \mid \alpha, \beta) = \begin{cases} \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, & 0 < x < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that $X \sim Be(\alpha, \beta)$, for some $\alpha > 0$ and $\beta > 0$. Then

$$E(X^r) = \frac{B(\alpha+r, \beta)}{B(\alpha, \beta)} = \frac{\Gamma\alpha+r\ \Gamma\alpha+\beta}{\Gamma\alpha\ \Gamma\alpha+\beta+r}, \quad r > -\alpha$$

$$\text{Mean} = \mu_1' = E(X) = \frac{\alpha}{\alpha+\beta}$$

$$\mu_2' = E(X^2) = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

$$\mu_2 = \sigma^2 = Var(X) = E(X^2) - (E(X))^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Mode $= M_o = \dfrac{\alpha-1}{\alpha+\beta-2}$, if $\alpha>1$ and $\alpha+\beta>2$

Skewness $= \beta_1 = \dfrac{\mu_3^2}{\mu_2^3} = \dfrac{2(\beta-\alpha)\sqrt{\alpha+\beta+1}}{\sqrt{\alpha\beta}\;(\alpha+\beta+2)}$

Kurtosis $= \gamma_1 = \dfrac{\mu_4}{\mu_2^2} = \dfrac{6\left[(\alpha-\beta)^2(\alpha+\beta+1)-\alpha\beta(\alpha+\beta+2)\right]}{\alpha\beta(\alpha+\beta+2)(\alpha+\beta+3)} + 3$

$= 6\dfrac{\left[\alpha^3+\alpha^2(1-2\beta)+\beta^2(1+\beta)-2\alpha\beta(2+\beta)\right]}{\alpha\beta(\alpha+\beta+2)(\alpha+\beta+3)}$

- Let $X \sim Be(\alpha, \alpha)$, $\alpha>0$. Then

$$f(x|\alpha,\beta) = \begin{cases} \dfrac{x^{\alpha-1}(1-x)^{\alpha-1}}{B(\alpha,\alpha)}, & 0<x<1 \\ 0, & \text{otherwise} \end{cases}$$

Clearly $X \stackrel{d}{=} 1-X$

$\Rightarrow X-\frac{1}{2} \stackrel{d}{=} \frac{1}{2}-X$

Thus if $X \sim Be(\alpha,\alpha)$ then the distribution of X is symmetric about $1/2$.

Theorem 5.3.1 (Relationship between Beta and Binomial Distribution). For $m, n \in \mathbb{N}$ and $x \in (0,1)$, let $X \sim Be(m,n)$ and $Y \sim Bin(m+n-1, x)$. Then

$$P(X \le x) = P(Y \ge m)$$

i.e.,
$$\frac{1}{B(m,n)}\int_0^x t^{m-1}(1-t)^{n-1}\,dt = \sum_{j=m}^{m+n-1}\binom{m+n-1}{j}x^j(1-x)^{m+n-1-j}$$

Proof Fix $m,n\in\mathbb{N}$ and $x\in(0,1)$. Let

$$I_{m,n} = \text{LHS} = \frac{\lfloor m+n-1}{\lfloor m-1\,\lfloor n-1}\int_0^x t^{m-1}(1-t)^{n-1}\,dt$$

$$= \binom{m+n-1}{m}x^m(1-x)^{n-1} + \frac{\lfloor m+n-1}{\lfloor m\,\lfloor n-2}\int_0^x t^m(1-t)^{n-2}\,dt$$

$$= \binom{m+n-1}{m}x^m(1-x)^{n-1} + I_{m+1,n-1}$$

M.g.f. $M_X(t) = E(e^{tX}) = \frac{1}{B(\alpha,\beta)} \int_0^1 e^{tx} x^{\alpha-1}(1-x)^{\beta-1}\,dx$

$$= \frac{1}{B(\alpha,\beta)} \int_0^1 \left(\sum_{j=0}^{\infty} \frac{t^j x^j}{\lfloor j} \right) x^{\alpha-1}(1-x)^{\beta-1}\,dx$$

$$= \frac{1}{B(\alpha,\beta)} \sum_{j=0}^{\infty} \frac{t^j}{\lfloor j} \int_0^1 x^{\alpha+j-1}(1-x)^{\beta-1}\,dx$$

$$= \frac{1}{B(\alpha,\beta)} \sum_{j=0}^{\infty} \frac{B(\alpha+j,\beta)}{\lfloor j} t^j, \qquad t \in \mathbb{R}.$$

Example 5.3.3.1 Time (in hours) to finish a job follows beta distribution with mean $\frac{1}{3}$ hrs. and variance $\frac{2}{63}$ hrs. Find the probability that the job will be finished in 30 minutes.

Proof Define

X = time to finish job (in hours)

$\sim Be(\alpha, \beta)$, say.

$E(X) = \frac{1}{3}$, $Var(X) = \frac{2}{63}$

$$\Rightarrow \frac{\alpha}{\alpha+\beta} = \frac{1}{3}, \quad \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{2}{63}$$

$\Rightarrow \alpha = 2$ and $\beta = 4$ $\Rightarrow X \sim Be(2, 4)$

Required probability $= P(X < \frac{1}{2}) = \frac{1}{B(2,4)} \int_0^{1/2} x(1-x)^3\,dx$

$$= \frac{13}{16}.$$

Theorem 5.3.3.2 (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 i.i.d. $Exp(\theta)$ r.v.s. Then $Y = \frac{X_1}{X_1+X_2} \sim U(0, 1)$.

Proof (a) The joint pdf of $\underset{\sim}{X} = (X_1, X_2)$ is

$$f_{\underset{\sim}{X}}(x_1, x_2) = \prod_{i=1}^{2} f_{X_i}(x_i) = \prod_{i=1}^{2} \left\{ \frac{1}{\Gamma\alpha_i \, \theta^{\alpha_i}} x_i^{\alpha_i - 1} e^{-\frac{x_i}{\theta}} I_{(0,\infty)}(x_i) \right\}$$

$$= \begin{cases} \dfrac{1}{\Gamma\alpha_1 \Gamma\alpha_2 \, \theta^{\alpha_1+\alpha_2}} x_1^{\alpha_1-1} x_2^{\alpha_2-1} e^{-\frac{x_1+x_2}{\theta}}, & \text{if } x_1 > 0, x_2 > 0 \\ 0, & \text{otherwise} \end{cases}.$$

Here $S_{\underset{\sim}{X}} = (0,\infty)^2$. Let

$h_1(x_1, x_2) = Y_1 = X_1 + X_2$ and $h_2(x_1, x_2) = Y_2 = \dfrac{X_1}{X_1+X_2}$

Then $\underset{\sim}{h} = (h_1, h_2) : S_{\underset{\sim}{X}} \to \mathbb{R}^2$ is 1-1 with inverse image (h_1^{-1}, h_2^{-1}),

where $h_1^{-1}(y_1, y_2) = y_1 y_2$, $h_2^{-1}(y_1, y_2) = y_1(1-y_2)$

$$J = \begin{vmatrix} y_2 & y_1 \\ 1-y_2 & -y_1 \end{vmatrix} = -y_1$$

$\underset{\sim}{h}^{-1}(\underset{\sim}{y}) \in S_{\underset{\sim}{X}} \iff y_1 y_2 > 0,\ y_1(1-y_2) > 0 \iff y_1 > 0,\ 0 < y_2 < 1$

$\Rightarrow \underset{\sim}{h}(S_{\underset{\sim}{X}}) = (0,\infty) \times (0,1)$.

Thus the joint p.d.f. of $\underset{\sim}{Y} = (Y_1, Y_2)$ is

$$f_{\underset{\sim}{Y}}(y_1, y_2) = \frac{1}{\Gamma\alpha_1 \Gamma\alpha_2 \, \theta^{\alpha_1+\alpha_2}} (y_1 y_2)^{\alpha_1-1} (y_1(1-y_2))^{\alpha_2-1} e^{-\frac{y_1 y_2 + y_1(1-y_2)}{\theta}} |-y_1| \, I_{(0,\infty)\times(0,1)}(y_1, y_2)$$

$$= \left\{ \frac{e^{-\frac{y_1}{\theta}} y_1^{\alpha_1+\alpha_2-1}}{\Gamma(\alpha_1+\alpha_2)} I_{(0,\infty)}(y_1) \right\} \left\{ \frac{1}{B(\alpha_1, \alpha_2)} y_2^{\alpha_1-1} (1-y_2)^{\alpha_2-1} I_{(0,1)}(y_2) \right\}$$

$$= f_{Y_1}(y_1) \, f_{Y_2}(y_2),$$

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.

5.3.4. Normal Distribution

Recall that

$$\sqrt{\pi} = \Gamma\tfrac{1}{2} = \int_0^\infty e^{-t} t^{-1/2}\, dt$$

$$= 2\int_0^\infty e^{-x^2}\, dx = \int_{-\infty}^\infty e^{-x^2}\, dx$$

$$= \frac{1}{\sqrt{2}} \int_{-\infty}^\infty e^{-\frac{t^2}{2}}\, dt$$

$$\Rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{t^2}{2}}\, dt$$

$$\Rightarrow \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{(t-\mu)^2}{2\sigma^2}}\, dt = 1 \quad \forall\ \mu \in \mathbb{R} \text{ and } \sigma > 0.$$

Definition 5.3.4.1 Let $\mu \in \mathbb{R}$ and $\sigma > 0$ be given constants. An absolutely continuous type r.v. X is said to follow a normal distribution with parameters $\mu \in \mathbb{R}$ and $\sigma > 0$ [written as $X \sim N(\mu, \sigma^2)$] if its p.d.f. is given by

$$f(x \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

The $N(0,1)$ distribution is called the standard normal distribution. The p.d.f. and d.f. of standard normal distribution are denoted by $\phi(z)$ and $\Phi(z)$, respectively, so that

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty, \qquad \Phi(z) = \int_{-\infty}^{z} \phi(t)\, dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\, dt, \quad z \in \mathbb{R}$$

- $X \sim N(\mu, \sigma^2) \Rightarrow f(\mu - x \mid \mu, \sigma) = f(\mu + x \mid \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \quad x \in \mathbb{R}$

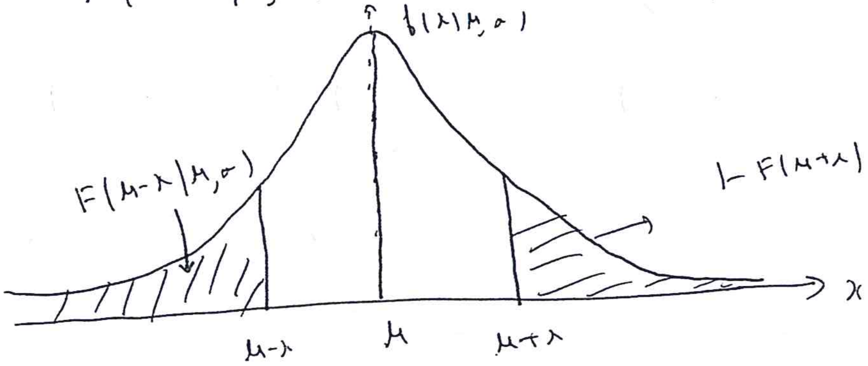
$$\Rightarrow X - \mu \overset{d}{=} \mu - X \quad (\text{Dist. of } X \text{ is symmetric about } \mu)$$

$$\Rightarrow E(X) = \mu \quad \text{and} \quad F(\mu \mid \mu, \sigma) = \tfrac{1}{2}.$$

Moreover

$$P(X-\mu \le x) = P(\mu - X \le x) \quad \forall x \in \mathbb{R}$$

$$\Rightarrow F(\mu + x \mid \mu, \sigma) = 1 - F(\mu - x \mid \mu, \sigma), \quad \forall x \in \mathbb{R}$$



In particular

$$\Phi(0) = \frac{1}{2} \quad \text{and} \quad \Phi(z) + \Phi(-z) = 1 \quad \forall z \in \mathbb{R}.$$

The pdf $f(x|\mu,\sigma)$ $\uparrow$ in $(-\infty, \mu)$ and $\downarrow$ in (μ, ∞)

$$\Rightarrow \text{Mode} = m_0 = \mu$$

Thus $\quad$ Mean = Median = Mode = μ.

<u>M.G.F.</u> $\quad$ Let $X \sim N(\mu, \sigma^2)$. Then

$$M_X(t) = E(e^{tX}) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\mu+\sigma z)} e^{-\frac{z^2}{2}} dz \quad \left(\frac{x-\mu}{\sigma} = z\right)$$

$$= e^{\mu t + \frac{\sigma^2 t^2}{2}} \times \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(z-t\sigma)^2}{2}} dz}_{=1}$$

$$= e^{\mu t + \frac{\sigma^2 t^2}{2}}, \quad t \in \mathbb{R}$$

- Let $X \sim N(\mu, \sigma^2)$. Then $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$.

$$\left[M_Z(t) = E\left(e^{\frac{t(X-\mu)}{\sigma}}\right) = e^{-\frac{\mu t}{\sigma}} M_X\left(\frac{t}{\sigma}\right) = e^{-\frac{\mu t}{\sigma}} e^{\frac{\mu t}{\sigma} + \frac{\sigma^2}{2} \times \frac{t^2}{\sigma^2}} \right.$$

$$= e^{t^2/2}, \quad \forall t \in \mathbb{R}$$

$\hookrightarrow$ m.g.f. of $N(0,1)$

$\Rightarrow \quad Z \sim N(0,1)$

$\boxed{\text{Theorem 3.3.4.1}}$ Let $X \sim N(\mu, \sigma^2)$ and let $Z \sim N(0,1)$

(a) for $a \neq 0$, $b \in \mathbb{R}$, $Y = aX + b \sim N(a\mu + b, \; a^2\sigma^2)$

(b) $Z \overset{d}{=} \frac{X-\mu}{\sigma} \sim N(0,1)$.

(c)
$$E(Z^r) = \begin{cases} 0, & \text{if } r = 1, 3, 5, \ldots \\ \dfrac{\lfloor r}{2^{\frac{r}{2}} \lfloor \frac{r}{2}}, & \text{if } r = 2, 4, 6, \ldots \end{cases}$$

(d) Mean $= \mu_1' = E(X) = \mu$;
Variance $= \mu_2 = \sigma^2$;
Coefficient of skewness $= \beta_1 = 0$;
Kurtosis $= \gamma_1 = 3$;

(e) $Z^2 \sim \chi^2_1$.

Proof. (a) $M_Y(t) = E(e^{t(aX+b)})$

$$= e^{bt} M_X(at)$$

$$= e^{bt} e^{\mu at + \frac{\sigma^2 a^2 t^2}{2}}$$

$$= e^{(a\mu + b)t + \frac{\sigma^2 a^2 t^2}{2}}, \quad t \in \mathbb{R}$$

$\longrightarrow$ m.g.f. of $N(a\mu + b, \; a^2\sigma^2)$

$\Rightarrow Y \sim N(a\mu + b, \; a^2\sigma^2)$

(b) Follows from (a) by taking $a = \frac{1}{\sigma}$ and $b = -\frac{\mu}{\sigma}$.

(c) $M_Z(t) = e^{t^2/2}, \quad t \in \mathbb{R}$

$$= \sum_{k=0}^{\infty} \frac{t^{2k}}{2^k \lfloor k}.$$

$E(Z^r) =$ Coefficient of $\frac{t^r}{\lfloor r}$ in the expansion of $M_X(t)$

$$= \begin{cases} 0, & \text{if } r = 1, 3, 5, \ldots \\ \dfrac{\lfloor r}{2^{\frac{r}{2}} \lfloor \frac{r}{2}}, & \text{if } r = 2, 4, 6, \ldots \end{cases}$$

(d) $Z \stackrel{d}{=} \frac{X-\mu}{\sigma}$

$E\left(\frac{X-\mu}{\sigma}\right) = E(Z) = 0 \quad \Rightarrow \quad \mu_1' = E(X) = \mu$

$E\left(\left(\frac{X-\mu}{\sigma}\right)^2\right) = E(Z^2) = 1 \quad \Rightarrow \quad \mu_2 = E((X-\mu)^2) = \sigma^2$

$\mu_3 = E\left(\left(\frac{X-\mu}{\sigma}\right)^3\right) = E(Z^3) = 0 \Rightarrow \mu_3 = E((X-\mu)^3) = 0$

$\mu_4 = E\left(\left(\frac{X-\mu}{\sigma}\right)^4\right) = E(Z^4) \Rightarrow \mu_4 = 3\sigma^4$
$= 3$

Coeff. of Skewness $= \beta_1 = \frac{\mu_3}{\mu_2^3} = 0$

Kurtosis $= \delta_1 = \frac{\mu_4}{\mu_2^2} = 3$

(e) Let $Y = Z^2$. Then

$$M_Y(t) = E(e^{tZ^2})$$
$$= \int_{-\infty}^{\infty} e^{tz^2} \frac{e^{-z^2/2}}{\sqrt{2\pi}} \, dz$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(1-2t)}{2} z^2} \, dz = (1-2t)^{-1/2}, \quad t < 1/2$$

$\downarrow$

mgf of χ_1^2

$\Rightarrow Z^2 \sim \chi_1^2$

$\boxed{\text{Corollary 5.3.4}}$ Let $X_1, \ldots, X_k$ be independent and let $X_i \sim N(\mu_i, \sigma_i^2)$
$-\infty < \mu_i < \infty$, $\sigma_i > 0$, $i = 1, \ldots, k$. Then

$$\sum_{i=1}^{k} \left(\frac{X_i - \mu_i}{\sigma_i}\right)^2 \sim \chi_k^2.$$

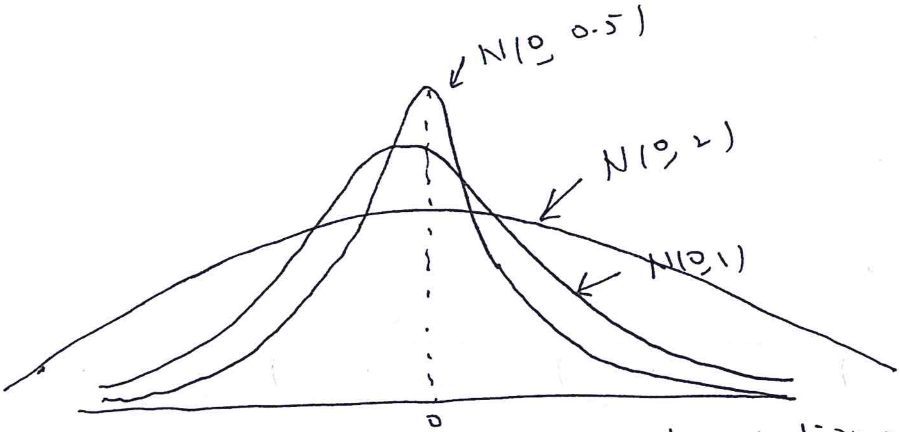


Figure: Pdfs of $N(0, \sigma^2)$ distributions

J.3.4.1

Remark (i) In $N(\mu, \sigma^2)$ distribution the parameters $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ are, respectively, the mean and the variance of the distribution.

(ii) If $X \sim N(\mu, \sigma^2)$ then $P(X \le x) = P\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$, $x \in \mathbb{R}$.

Let τ_α be the $(1-\alpha)$-th quantile of Φ ($N(0,1)$ distribution):

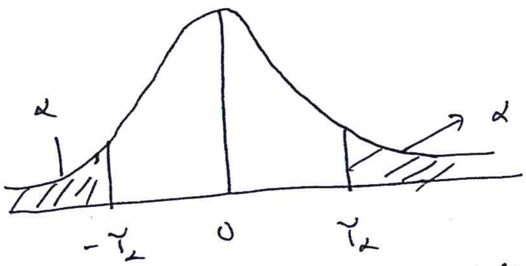


Figure: $(1-\alpha)$-th quantile of $N(0,1)$ distribution

Then

$$\Phi(-\tau_\alpha) = 1 - \Phi(\tau_\alpha) = \alpha$$

α	0.001	0.05	0.01	0.025	0.05	0.1	0.25
τ_α	3.092	2.5758	2.326	1.96	1.6499	1.282	0.675

Table: $(1-\alpha)$-th quantiles of $N(0,1)$ distribution

Tables for values of $\Phi(x)$ for different values of x are available in various text books

Example 5.3.4.1 Let $X \sim N(2, 4)$. Find $P(X \le 0)$, $P(|X| \ge 2)$, $P(1 < X \le 3)$ and $P(X \le 3 \mid X > 1)$.

Solution $P(X \le 0) = \Phi\left(\frac{0-2}{2}\right) = \Phi(-1) = 1 - \Phi(1) = .1587$

$$P(|X| \ge 2) = P(X \le -2) + P(X \ge 2)$$
$$= \Phi\left(\frac{-2-2}{2}\right) + 1 - \Phi\left(\frac{2-2}{2}\right)$$
$$= \Phi(-2) + 1 - \Phi(0)$$
$$= 0.0228 + 0.5 = 0.5228$$

$$P(1 < X \le 3) = P(X \le 3) - P(X \le 1)$$
$$= \Phi\left(\frac{3-2}{2}\right) - \Phi\left(\frac{1-2}{2}\right) = \Phi(0.5) - \Phi(-0.5)$$
$$= 2\Phi(0.5) - 1 = 2 \times .6915 - 1 = .383$$

$$P(X \le 3 \mid X > 1) = \frac{P(1 < X \le 3)}{P(X > 1)} = \frac{0.383}{1 - \Phi\left(\frac{1-2}{2}\right)} = \frac{0.383}{\Phi(0.5)}$$
$$= 0.55599$$

Theorem 5.3.4.2 Let $X_1, \ldots, X_k$ be independent rvs and let $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2, \ldots k$. Let $a_1, \ldots, a_k$ be real constants such that $\sum_{i=1}^{k} a_i^2 > 0$. Then

$$Y = \sum_{i=1}^{k} a_i X_i \sim N\left(\sum_{i=1}^{k} a_i \mu_i, \sum_{i=1}^{k} a_i^2 \sigma_i^2\right).$$

Proof
$$M_Y(t) = E\left(e^{t \sum_{i=1}^{k} a_i X_i}\right)$$
$$= E\left(\prod_{i=1}^{k} e^{t a_i X_i}\right) = \prod_{i=1}^{k} E\left(e^{t a_i X_i}\right) \quad (\text{independence})$$
$$= \prod_{i=1}^{k} M_{X_i}(t a_i)$$
$$= \prod_{i=1}^{k} e^{\mu_i t a_i + \frac{\sigma_i^2 t^2 a_i^2}{2}} = e^{\left(\sum_{i=1}^{k} a_i \mu_i\right) t + \frac{\left(\sum_{i=1}^{k} a_i^2 \sigma_i^2\right) t^2}{2}}$$
$$\longrightarrow \text{m.g.f. of } N\left(\sum_{i=1}^{k} a_i \mu_i, \sum_{i=1}^{k} a_i^2 \sigma_i^2\right)$$

By uniqueness of m.g.f.'s $Y \sim N\left(\sum_{i=1}^{k} a_i \mu_i, \sum_{i=1}^{k} a_i^2 \sigma_i^2\right)$.

Theorem 5.3.4.3 Let $X_1, \ldots, X_n$ $(n \geq 2)$ 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) $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$;

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

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

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

Then

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

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

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

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

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

Then $\sum_{i=1}^{n} u_i = t_{n+1}$ and

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

$$M_{\underset{\sim}{Y},\bar{X}}(\underset{\sim}{t}) = E\left(e^{\sum_{i=1}^{n} u_i X_i}\right)$$

$$= \prod_{i=1}^{n} M_{X_i}(u_i)$$

$$= \prod_{i=1}^{n} e^{\mu u_i + \frac{\sigma^2 u_i^2}{2}}$$

$$= e^{\mu \sum_{i=1}^{n} u_i + \frac{\sigma^2}{2} \sum_{i=1}^{n} u_i^2}$$

$$= e^{\mu t_{n+1} + \frac{\sigma^2}{2} \left\{ \sum_{i=1}^{n} (t_i-\bar{t})^2 + \frac{t_{n+1}^2}{n} \right\}}$$

$$= \left\{ e^{\mu t_{n+1} + \frac{\sigma^2 t_{n+1}^2}{2n}} \right\} \times \left\{ e^{\frac{\sigma^2}{2} \sum_{i=1}^{n} (t_i-\bar{t})^2} \right\}$$

$$M_{\underset{\sim}{Y}}(t_1,\dots,t_n) = M_{\underset{\sim}{Y},\bar{X}}(t_1,\dots,t_n,0) = e^{\frac{\sigma^2}{2} \sum_{i=1}^{n} (t_i-\bar{t})^2}, \quad (t_1,\dots,t_n) \in \mathbb{R}^n$$

$$M_{\bar{X}}(t_{n+1}) = M_{\underset{\sim}{Y},\bar{X}}(0,\dots,0,t_{n+1}) = e^{\mu t_{n+1} + \frac{\sigma^2 t_{n+1}^2}{2n}}, \quad t_{n+1} \in \mathbb{R}$$

$$\Rightarrow \quad M_{\underset{\sim}{Y},\bar{X}}(\underset{\sim}{t}) = M_{\underset{\sim}{Y}}(t_1,\dots,t_n)\, M_{\bar{X}}(t_{n+1}), \quad \forall\ \underset{\sim}{t} = (t_1,\dots,t_n,t_{n+1}) \in \mathbb{R}^{n+1}$$

$\Rightarrow$ $\underset{\sim}{Y}$ and $\bar{X}$ are independent

$\Rightarrow$ $\sum_{i=1}^{n} (X_i-\bar{X})^2$ and $\bar{X}$ are independent.

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

Let

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

Then, by (ii), W and T are independent r.v.s. Also

$W \sim \chi^2_1$ and $V = \sum_{i=1} Z_i^2 \sim \chi^2_n$.

$$V = \sum_{i=1}^{n} Z_i^2 = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2$$

$$= \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \bar{X} + \bar{X} - \mu)$$

$$= \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \bar{X})^2 + \frac{n(\bar{X}-\mu)^2}{\sigma^2}$$

$$= \frac{(n-1)S^2}{\sigma^2} + \frac{n(\bar{X}-\mu)^2}{\sigma^2}$$

$$= T + W$$

$$\Rightarrow M_V(t) = M_T(t)\, M_W(t)$$

$$\Rightarrow M_T(t) = \frac{M_V(t)}{M_W(t)} = \frac{(1-2t)^{-\frac{n}{2}}}{(1-2t)^{-\frac{1}{2}}}, \quad t < \frac{1}{2}$$

$$= (1-2t)^{-\frac{n-1}{2}}, \quad t < \frac{1}{2}$$

$\longrightarrow$ m.g.f. of χ^2_{n-1}

$$\Rightarrow T = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

(iv) $T = \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{\nu}$, where $\nu = n-1$.

Thus

$$E(T^{r}) = \int_0^\infty t^{r} \frac{1}{2^{\nu/2}\Gamma\frac{\nu}{2}} e^{-t/2}\, t^{\nu/2-1}\, dt$$

$$= \frac{1}{2^{\nu/2}\Gamma\frac{\nu}{2}} \int_0^\infty e^{-t/2}\, t^{\frac{\nu+2r}{2}-1}\, dt$$

$$\Rightarrow \quad = \frac{2^{\frac{\nu+2r}{2}}\, \Gamma\frac{\nu+2r}{2}}{2^{\nu/2}\,\Gamma\frac{\nu}{2}} = 2^{r} \frac{\Gamma\frac{\nu+2r}{2}}{\Gamma\frac{\nu}{2}}, \quad r > -\frac{\nu}{2}$$

$$\Rightarrow \frac{(n-1)^{r}}{\sigma^{2r}} E(S^{2r}) = 2^{r} \frac{\Gamma\frac{\nu+2r}{2}}{\Gamma\frac{\nu}{2}}$$

$$\Rightarrow E(S^r) = \left(\frac{2}{n-1}\right)^{\frac{r}{2}} \frac{\Gamma\left(\frac{\nu+r}{2}\right)}{\Gamma\left(\nu/2\right)} \sigma^r, \quad r>0$$

$$\Rightarrow \quad E(S^r) = \left(\frac{2}{n-1}\right)^{r/2} \frac{\Gamma\left(\frac{n-1+r}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \sigma^r, \quad r>0$$

$$E(S) = \sqrt{\frac{2}{n-1}} \frac{\Gamma\left(\frac{n}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \sigma$$

$$E(S^2) = \frac{2}{n-1} \frac{\Gamma\left(\frac{n-1}{2}+1\right)}{\Gamma\left(\frac{n-1}{2}\right)} \sigma^2 = \sigma^2$$

$$E(S^4) = \left(\frac{2}{n-1}\right)^2 \frac{\Gamma\left(\frac{n-1}{2}+2\right)}{\Gamma\left(\frac{n-1}{2}\right)} \sigma^4 = \frac{n+1}{n-1}\sigma^4$$

$$Var(S^2) = E(S^4) - (E(S^2))^2 = \frac{2\sigma^4}{n-1}.$$

Remark 5.3.4.2 Let $X_1, \ldots, X_n$ 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 $Var(X_1) = \sigma^2$. Then

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \mu. \quad Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{\sigma^2}{n}.$$

$$E[(n-1)S^2] = E\left[\sum_{i=1}^{n}(X_i-\bar{X})^2\right]$$

$$\Rightarrow (n-1)E(S^2) = E\left[\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right]$$

$$= \sum_{i=1}^{n} E(X_i^2) - nE(\bar{X}^2)$$

$$= n\left[E(X_1^2) - E(\bar{X}^2)\right]$$

$$= n\left[Var(X_1) + (E(X_1))^2 - Var(\bar{X}) - (E(\bar{X}))^2\right]$$

$$= n\left[\sigma^2 + \mu^2 - \frac{\sigma^2}{n} - \mu^2\right] = (n-1)\sigma^2$$

$$\Rightarrow E(S^2) = \sigma^2$$

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.

5.3.5 Distributions Based on Sampling From Normal Distribution

Definition 5.3.5.1 (a) For a positive integer m a rv X is said to have the Student t-distribution with m degrees of freedom (written as $X \sim t_m$) if the p.d.f. of X is given by

$$f(x|m) = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{m\pi}\,\Gamma\left(\frac{m}{2}\right)} \frac{1}{\left(1+\frac{x^2}{m}\right)^{\frac{m+1}{2}}}, \quad -\infty < x < \infty$$

(b) For positive integers n_1 and n_2 a rv X is said to have the Snedcor F distribution with (n_1, n_2) degrees of freedom (written as $X \sim F_{n_1, n_2}$) if its p.d.f is given by

$$f(x|n_1,n_2) = \begin{cases} \dfrac{\frac{n_1}{n_2}}{B\left(\frac{n_1}{2}, \frac{n_2}{2}\right)} \dfrac{\left(\frac{n_1}{n_2}x\right)^{\frac{n_1}{2}-1}}{\left(1+\frac{n_1}{n_2}x\right)^{\frac{n_1+n_2}{2}}}, & \text{if } x>0 \\ 0, & \text{otherwise} \end{cases}$$

Remark 5.3.5.1 (a) $X \sim t_m \Rightarrow f(x|m) = f(-x|m), \ \forall x$

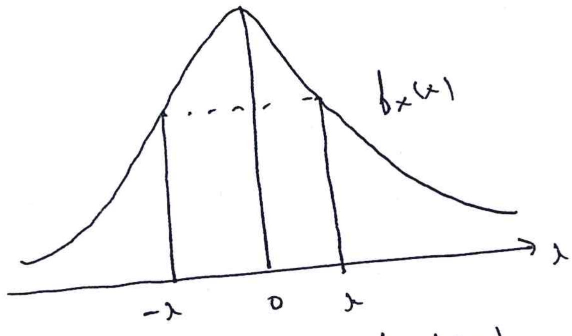
$\Rightarrow X \stackrel{d}{=} -X$

$\Rightarrow$ distribution of X is symmetric about 0

$\Rightarrow M_e = 0$ and $E(X)=0$, provided it exists

(b) $X \sim t_m \Rightarrow$ pdf $f(x|m) \uparrow$ in $(-\infty, 0)$, $\downarrow (0, \infty)$

$\Rightarrow M_0 = 0$



<u>Figure</u> pdf of $f(x|m)$

(c) t_1 distribution is nothing but Cauchy distribution with pdf

$$f(x|1) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}, \quad -\infty < x < \infty$$

$\Rightarrow E(X)$ does not exist.

(c) $X \sim F_{n_1, n_2} \Rightarrow f(x|n_1, n_2) = \begin{cases} \dfrac{\frac{n_1}{n_2}}{B(\frac{n_1}{2}, \frac{n_2}{2})} \left(\dfrac{\frac{n_1}{n_2}x}{1+\frac{n_1}{n_2}x}\right)^{\frac{n_1}{2}-1} \left(1 - \dfrac{\frac{n_1}{n_2}x}{1+\frac{n_1}{n_2}x}\right)^{n_2-1} \times \dfrac{1}{(1+\frac{n_1}{n_2}x)^2}, & \text{if } x > 0 \\ 0, & \text{Otherwise} \end{cases}$

$$\Rightarrow Y = \frac{\frac{n_1}{n_2}X}{1+\frac{n_1}{n_2}X} \sim Be\left(\frac{n_1}{2}, \frac{n_2}{2}\right)$$

$\boxed{\text{Theorem } 5.3.5.[illegible]}$ (a). Let $Z \sim N(0,1)$ and let $Y \sim \chi^2_m$, $m \in \{1, 2, \ldots\}$, be independent rvs. Then

$$T = \frac{Z}{\sqrt{\frac{Y}{m}}} \sim t_m.$$

(b) For positive integers n_1 and n_2 let $X_1 \sim \chi^2_{n_1}$ and $X_2 \sim \chi^2_{n_2}$ be independent rvs. Then

$$U = \frac{X_1/n_1}{X_2/n_2} \sim F_{n_1, n_2}.$$

(c) Let $X \sim t_m$. Then $E(X^r)$ is not finite if $r \in \{m, m+1, \ldots\}$.

For $r \in \{1, 2, \ldots, m-1\}$ $(m \geq r+1)$

$$E(X^r) = \begin{cases} 0, & \text{if } r \text{ is odd} \\ \dfrac{m^{\frac{r}{2}} \, r! \, \Gamma\left(\frac{m-r}{2}\right)}{2^r \, \left(\frac{r}{2}\right)! \, \Gamma\left(\frac{m}{2}\right)}, & \text{if } r \text{ is even} \end{cases}$$

(d) If $X \sim t_m$ then

$$\text{mean} = \mu_1' = E(X) = 0, \quad m = 2, 3, \ldots$$

$$\text{Variance} = \mu_2 = E((X-\mu_1')^2) = \frac{m}{m-2}, \quad m = 3, 4, \ldots$$

$$\text{Coefficient of skewness} = \beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0, \quad m = 4, 5, 6, \ldots$$

$$\gamma_1 = \text{Kurtosis} = \frac{3(m-2)}{m-4}, \quad m = 5, 6, \ldots$$

(e) Let n_1, n_2 and r be positive integers, and let $X \sim F_{n_1, n_2}$. Then, for $n_2 \in \{1, 2, \ldots, 2r\}$ and $r \geq \frac{n_2}{2}$, $E(X^r)$ is not finite. For $n_2 \in \{2r+1, 2r+2, \ldots\}$, $r > \frac{n_2 - 1}{2}$

$$E(X^r) = \left(\frac{n_2}{n_1}\right)^r \prod_{i=1}^{r} \left(\frac{n_1 + 2(i-1)}{n_2 - 2i}\right).$$

(f) If $X \sim F_{n_1, n_2}$ then

$$\text{mean} = \mu_1' = E(X) = \frac{n_2}{n_2 - 2}, \quad n_2 \in \{3, 4, \ldots\}$$

$$\text{Variance} = \mu_2 = E((X-\mu_1')^2) = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)}, \quad n_2 = 5, 6, \ldots$$

$$\text{Coefficient of skewness} = \beta_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2(2n_1+n_2-2)}{n_2-6}\sqrt{\frac{2(n_2-4)}{n_1(n_1+n_2-2)}},$$

$$n_2 \in \{7, 8, \ldots\}$$

$$\text{Kurtosis} = \gamma_1 = \frac{12\left[(n_2-2)^2(n_2-4) + n_1(n_1+n_2-2)(5n_2-22)\right]}{n_1(n_2-6)(n_2-8)(n_1+n_2-2)}$$

Proof (a) The joint pdf of (Y, Z) is given by

$$f_{Y,Z}(y, z) = f_Y(y) f_Z(z) = \begin{cases} \dfrac{1}{2^{\frac{m+1}{2}} \Gamma\left(\frac{m}{2}\right) \sqrt{\pi}} e^{-\frac{(y+z^2)}{2}} y^{\frac{m}{2}-1}, & y > 0, -\infty < z < \infty \\ 0 & \text{otherwise} \end{cases}$$

Let $U = \sqrt{\frac{Y}{m}}$

$S_{Y,Z} = (0, \infty) \times \mathbb{R}$. Let $\underset{\sim}{h} = (h_1, h_2): (0, \infty) \times \mathbb{R} \to \mathbb{R}^2$, where

$h_1(y, z) = \dfrac{z}{\sqrt{\frac{y}{m}}}$ and $h_2(y, z) = \sqrt{\frac{y}{m}}$.

The transformation $\underset{\sim}{h}: S_{Y,Z} \to \mathbb{R}$ is 1-1 with inverse transformation

$\underset{\sim}{h}^{-1} = (h_1^{-1}, h_2^{-1})$, where

$h_1^{-1}(t,u) = mu^2$, $\quad h_2^{-1}(t,u) = tu$

$$J = \begin{vmatrix} 0 & 2mu \\ u & t \end{vmatrix} = -2mu^2$$

$\underset{\sim}{h}(S_{YZ}) = \{(t,u) : mu^2 > 0, \; -\infty < tu < \infty\}$

$= \{(t,u) : u > 0, \; t \in \mathbb{R}\} = \mathbb{R} \times (0,\infty)$

The jt pdf of (T,U) is given by

$f_{T,U}(t,u) = f_{Y,Z}(h_1^{-1}(t,u), h_2^{-1}(t,u)) \, |J| \, I_{h(S_{YZ})}(t,u)$

$$= \frac{1}{2^{\frac{m+1}{2}} \Gamma\frac{m}{2} \sqrt{\pi}} e^{-\frac{(mu^2 + t^2u^2)}{2}} (mu^2)^{\frac{m}{2}-1} |-2mu^2| \, I_{\mathbb{R}\times(0,\infty)}(t,u)$$

$$= \frac{m^{\frac{m}{2}}}{2^{\frac{m-1}{2}} \Gamma\frac{m}{2} \sqrt{\pi}} u^m e^{-\frac{(m+t^2)u^2}{2}} I_{\mathbb{R}}(t) \, I_{(0,\infty)}(u).$$

The marginal p.d.f. of T is

$$f_T(t) = \int_{-\infty}^{\infty} f_{T,U}(t,u)$$

$$= \frac{m^{\frac{m}{2}}}{2^{\frac{m-1}{2}} \Gamma\frac{m}{2} \sqrt{\pi}} \int_0^{\infty} u^m e^{-\frac{(m+t^2)}{2}u^2} du$$

$$= \frac{1}{\sqrt{m\pi}\,\Gamma\frac{m}{2}} \cdot \frac{1}{\left(1+\frac{t^2}{m}\right)^{\frac{m+1}{2}}} \int_0^{\infty} y^{\frac{m-1}{2}} e^{-y} dy$$

$(u^2 = y)$

$$= \frac{\Gamma\frac{m+1}{2}}{\sqrt{m\pi}\,\Gamma\frac{m}{2}} \frac{1}{\left(1+\frac{t^2}{m}\right)^{\frac{m+1}{2}}}, \quad t \in \mathbb{R}$$

$\longrightarrow$ p.d.f of t_m

(b) The joint p.d.f. of $\underset{\sim}{X} = (X_1, X_2)$ is given by

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

$$= \frac{1}{2^{\frac{n_1+n_2}{2}} \Gamma\frac{n_1}{2} \Gamma\frac{n_2}{2}} e^{-\frac{x_1+x_2}{2}} x_1^{\frac{n_1}{2}-1} x_2^{\frac{n_2}{2}-1} I_{(0,\infty)\times(0,\infty)}(x_1, x_2)$$

Let $V = \frac{X_2}{n_2}$. $S_{\underset{\sim}{X}} = (0,\infty)\times(0,\infty)$ Consider the transformation $\underset{\sim}{h} = (h_1, h_2) : (0,\infty)\times(0,\infty) \to \mathbb{R}$ defined by

$$h_1(x_1, x_2) = \frac{x_1/n_1}{x_2/n_2} \quad \text{and} \quad h_2(x_1, x_2) = \frac{x_2}{n_2}$$

so that $U = h_1(X_1, X_2)$ and $V = h_2(X_1, X_2)$.

The transformation $\underset{\sim}{h} : (0,\infty)\times(0,\infty) \to \mathbb{R}^2$ is 1-1 with inverse transformation $\underset{\sim}{h}^{-1} = (h_1^{-1}, h_2^{-1})$, where

$$h_1^{-1}(u, v) = n_1 u v \;;\quad h_2(u, v) = n_2 v$$

$$J = \begin{vmatrix} n_1 v & n_1 u \\ 0 & n_2 \end{vmatrix} = n_1 n_2 v$$

$$\underset{\sim}{h}(S_{\underset{\sim}{X}}) = \{(u, v) : n_1 u v > 0,\ n_2 v > 0\} = \{(u, v) : u > 0, v > 0\} = (0,\infty)\times(0,\infty)$$

Thus the joint p.d.f. of (U, V) is

$$f_{U,V}(u, v) = f_{X_1, X_2}(h_1^{-1}(u, v), h_2^{-1}(u, v))\, |J|\, I_{\underset{\sim}{h}(S_{\underset{\sim}{X}})}(u, v)$$

$$= \frac{n_1^{\frac{n_1}{2}} n_2^{\frac{n_2}{2}}}{2^{\frac{n_1+n_2}{2}} \Gamma\frac{n_1}{2} \Gamma\frac{n_2}{2}} u^{\frac{n_1}{2}-1} v^{\frac{n_1+n_2}{2}-1} e^{-\frac{(n_2+n_1 u)v}{2}} I_{(0,\infty)}(u)\, I_{(0,\infty)}(v)$$

The p.d.f. of U is given by

$$f_U(u) = \int_{-\infty}^{\infty} f_{U,V}(u, v)\, dv$$

$$= \frac{n_1^{\frac{n_1}{2}} n_2^{\frac{n_2}{2}}}{2^{\frac{n_1+n_2}{2}} \Gamma\frac{n_1}{2} \Gamma\frac{n_2}{2}} u^{\frac{n_1}{2}-1} \int_0^{\infty} v^{\frac{n_1+n_2}{2}-1} e^{-\frac{(n_2+n_1 u)v}{2}}\, dv$$

$$= \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{2^{\frac{n_1+n_2}{2}} \Gamma\left(\frac{n_1}{2}\right) \Gamma\left(\frac{n_2}{2}\right)} \frac{\left(\frac{n_1}{n_2} u\right)^{\frac{n_1}{2}-1}}{\left(1+\frac{n_1}{n_2} u\right)^{\frac{n_1+n_2}{2}}} I_{(0,\infty)}(u)$$

$\rightarrow$ p.d.f. of F_{n_1, n_2} distribution

(c) Fix $m \in \{1, 2, \ldots\}$. Then

$$X \stackrel{d}{=} \frac{Z}{\sqrt{\frac{Y}{m}}},$$

where $Z \sim N(0,1)$ and $Y \sim \chi^2_m$ are independent

$$\Rightarrow E(X^r) = E\left(\left(\frac{Z}{\sqrt{\frac{Y}{m}}}\right)^r\right) \underset{\wedge}{=} \overset{= m^{r/2} E(Z^r Y^{-r/2})}{E(Z^r)\, m^{r/2}\, E(Y^{-r/2})} \quad \left(\begin{array}{l} Y \text{ and } Z \\ \text{are indept} \end{array}\right)$$

$$E(Z^r) = \begin{cases} 0, & \text{if } r = 1, 3, 5, \ldots \\ \dfrac{\lfloor r}{2^{r/2} \lfloor \frac{r}{2}}, & \text{if } r = 2, 4, 6, \ldots \end{cases}$$

$$E\left(Y^{-\frac{r}{2}}\right) = \frac{1}{2^{\frac{m}{2}} \Gamma\left(\frac{m}{2}\right)} \int_0^\infty y^{\frac{m-r}{2}-1} e^{-y/2} \, dy = \infty \quad \text{if } r \geq m$$

For $r < m$

$$E\left(Y^{-\frac{r}{2}}\right) = \frac{2^{\frac{m-r}{2}} \Gamma\left(\frac{m-r}{2}\right)}{2^{m/2} \Gamma\left(m/2\right)} = \frac{\Gamma\left(\frac{m-r}{2}\right)}{2^{r/2} \Gamma\left(m/2\right)}$$

$$\Rightarrow E(X^r) = \begin{cases} 0, & r \text{ is odd and } r < m \\ \dfrac{m^{\frac{r}{2}} \lfloor r \, \Gamma\left(\frac{m-r}{2}\right)}{2^r \lfloor \frac{r}{2} \, \Gamma\left(\frac{m}{2}\right)}, & r \text{ is even and } r < m \end{cases}$$

(d) Exercise

(e) Fix $n_1, n_2 \in \mathbb{N}$. Then

$$X \stackrel{d}{=} \frac{X_1/n_1}{X_2/n_2} = \frac{n_2}{n_1} \frac{X_1}{X_2},$$

where $X_1 \sim \chi^2_{n_1}$ and $X_2 \sim \chi^2_{n_2}$ are independent.

For $r \in \mathbb{N}$

$$E(X^r) = \left(\frac{n_2}{n_1}\right)^r E\left(\frac{X_1^r}{X_2^r}\right) = \left(\frac{n_2}{n_1}\right)^r E(X_1^r)\, E\left(\frac{1}{X_2^r}\right)$$

$$E(X_1^r) = \frac{1}{2^{\frac{n_1}{2}}\Gamma\frac{n_1}{2}} \int_0^\infty x^{\frac{n_1+2r}{2}-1} e^{-\frac{x}{2}}\,dx$$

$$= \frac{2^{\frac{n_1+2r}{2}}\Gamma\frac{n_1+2r}{2}}{2^{\frac{n_1}{2}}\Gamma\frac{n_1}{2}} = \prod_{i=1}^{r}(n_1+2(i-1)),\quad r \in \{1,2,\dots\}$$

$$E\left(\frac{1}{X_2^r}\right) = \begin{cases} \dfrac{2^{\frac{n_2-2r}{2}}\Gamma\frac{n_2-2r}{2}}{2^{\frac{n_2}{2}}\Gamma\frac{n_2}{2}}, & \text{if } n_2 > 2r \\ \infty & \text{if } n_2 \le 2r \end{cases}$$

$$= \begin{cases} \prod_{i=1}^{r}(n_2-2i), & \text{if } n_2 > 2r \\ \infty, & \text{if } n_2 \le 2r \end{cases}$$

$$\Rightarrow E(X^r) = \begin{cases} \left(\frac{n_2}{n_1}\right)^r \prod_{i=1}^{r}\left(\frac{n_1+2(i-1)}{n_2-2i}\right), & n_2 > 2r \\ \infty, & \text{if } n_2 \le 2r. \end{cases}$$

(f) Exercise

Corollary 5.3.5.1 Let $X_1, \dots, X_n$ $(n \ge 2)$ be a random sample from $N(\mu, \sigma^2)$ distribution, where $\mu \in \mathbb{R}$ 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

$$\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim t_{n-1}$$

__Proof.__ We know that

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$
$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$
independent

$$\frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0,1)$$
$$\Rightarrow \quad \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$
independent

$$\Rightarrow \quad \frac{\sqrt{n}(\bar{X}-\mu)/\sigma}{\sqrt{(n-1)S^2/\sigma^2}} \sim t_{n-1}$$

i.e $$\frac{\sqrt{n}(\bar{X}-\mu)}{S} \sim t_{n-1}.$$

Corollary 5.3.5.2 Let $X_1 \ldots X_m$ $(m \geq 2)$ and $Y_1 \ldots Y_n$ $(n \geq 2)$ be independent random samples (i.e. $\underset{\sim}{X} = (X_1 \cdots X_m)$ and $\underset{\sim}{Y} = (Y_1 \cdots Y_n)$ are independent) from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ distributions, respectively, where $\mu_i \in \mathbb{R}$, $i=1,2$, and $\sigma_i > 0$, $i=1,2$. Let $\bar{X} = \frac{1}{m}\sum_{i=1}^{m} X_i$, $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$, $S_1^2 = \frac{1}{m-1}\sum_{i=1}^{m}(X_i-\bar{X})^2$ and

$$S_2^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2.$$ Then

(a) $$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \sim N(0,1)$$

(b) $$\frac{1}{\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}} \; \frac{(\bar{X}-\bar{Y}-(\mu_1-\mu_2))}{\sqrt{\frac{(m-1)S_1^2/\sigma_1^2+(n-1)S_2^2/\sigma_2}{m+n-2}}} \sim t_{m+n-2}$$

(c) $$\frac{\sigma_2^2 S_1^2}{\sigma_1^2 S_1^2} \sim F_{m-1,\, n-1}$$

Proof $\bar{X} \sim N(\mu_1, \frac{\sigma_1^2}{m})$, $\bar{Y} \sim N(\mu_2, \frac{\sigma_2^2}{n})$, $\frac{(m-1)S_1^2}{\sigma_1^2} \sim \chi^2_{m-1}$ and $\frac{(n-1)S_2^2}{\sigma_2^2}$ are independent r.v.s. Thus

$$\bar{X} - \bar{Y} \sim N\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}\right)$$

$$\frac{(m-1)S_1^2}{\sigma_1^2} + \frac{(n-1)S_2^2}{\sigma_2^2} \sim t_{m+n-2}$$

$$\Rightarrow \quad \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0,1)$$

and

$$\frac{\left(\bar{X} - \bar{Y} - (\mu_1 - \mu_2)\right) \Big/ \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}{\sqrt{\left(\frac{(m-1)S_1^2}{\sigma_1^2} + \frac{(n-1)S_2^2}{\sigma_2^2}\right) \Big/ m+n-2}} \sim t_{m+n-2}.$$

Remark 5.3.5.1 (a) $X \sim t_m$

$$\Rightarrow \quad X \overset{d}{=} \frac{N(0,1)}{\sqrt{\chi^2_m / m}} \quad \text{independent}$$

$$\chi^2_1$$

$$\Rightarrow \quad X^2 \overset{d}{=} \frac{(N(0,1))^2}{\chi^2_m / m} \quad \text{independent}$$

$$= \frac{\chi^2_1 / 1}{\chi^2_m / m} \quad \text{independent}$$

$$\overset{d}{=} F_{1,m}$$

Thus

$$X \sim t_m \Rightarrow X^2 \sim F_{1,m}$$

(b) $X \sim F_{n_1, n_2} \Rightarrow X \stackrel{d}{=} \dfrac{\chi^2_{n_1}/n_1}{\chi^2_{n_2}/n_2}$ > independent

$$\Rightarrow \frac{1}{X} \stackrel{d}{=} \frac{\chi^2_{n_2}/n_2}{\chi^2_{n_1}/n_1} \text{ > independent}$$

$$\stackrel{d}{=} F_{n_2, n_1}$$

Thus $X \sim F_{n_1, n_2} \Rightarrow \dfrac{1}{X} \sim F_{n_2, n_1}$

(c) $X \sim t_m \Rightarrow$ Kurtosis $= \gamma_1 = \dfrac{3(m-2)}{m-4}$, $m > 4$

$\Rightarrow$ t_m distribution $(m > 4)$ is symmetric and leptokurtic (i.e. it has sharper peak and longer fatter tails compared to $N(0,1)$ distribution)

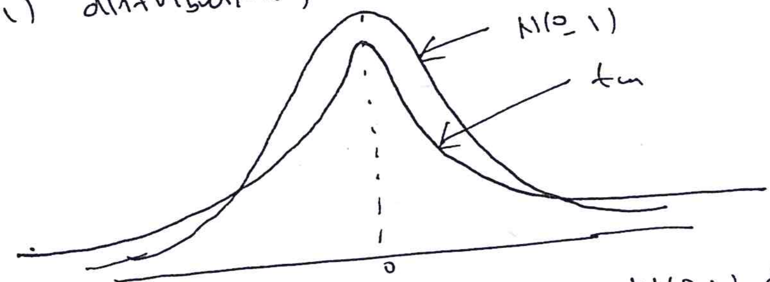


Figure: Pdfs of t_m and $N(0,1)$ distribution.

As $m \to \infty$, $\gamma_1 \to 0$. This suggests that for large d.f. m, t_m distribution behaves like $N(0,1)$ distribution.

(d) For various values of $m \in \mathbb{N}$ and $\alpha \in (0,1)$ the d.f. of t_m is tabulated in various text books.

(e) For fixed $n_1 \in \mathbb{N}$, $n_2 \in \mathbb{N}$ and $\alpha \in (0,1)$ let $f_{n_1, n_2, \alpha}$ be the $(1-\alpha)$-th quantile of $X \sim F_{n_1, n_2}$. Then

$$P(X \le f_{n_1, n_2, \alpha}) = 1 - \alpha$$

$$\Rightarrow P\left(\frac{1}{X} \le \frac{1}{f_{n_1, n_2, \alpha}}\right) = \alpha$$

$$\Rightarrow f_{n_2, n_1, 1-\alpha} = \frac{1}{f_{n_1, n_2, \alpha}} \quad \left(\text{since } \frac{1}{X} \sim F_{n_2, n_1}\right)$$

Example 5.3.5.1 Let $X_1, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$ distribution, where $\mu \in \mathbb{R}$, $\sigma > 0$ and $n \geq 2$. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ be the sample mean and sample variance, respectively. Evaluate $E\left(\frac{\bar{X}}{S}\right)$, for $n > 2$.

Solution

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$
$$Y = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$
independent

$$\Rightarrow E\left(\frac{\bar{X}}{S}\right) = \frac{\sqrt{n-1}}{\sigma} E\left(\bar{X} Y^{-\frac{1}{2}}\right)$$

$$= \frac{\sqrt{n-1}}{\sigma} E(\bar{X})\, E\left(Y^{-\frac{1}{2}}\right) \qquad \text{(independence)}$$

$$= \frac{\sqrt{n-1}}{\sigma}\, \mu \int_0^\infty \frac{e^{-\frac{y}{2}}\, y^{\frac{n-2}{2}-1}}{2^{\frac{n-1}{2}}\, \Gamma\left(\frac{n-1}{2}\right)}\, dy$$

$$= \frac{\sqrt{n-1}}{\sigma}\, \mu\, \frac{2^{\frac{n-2}{2}}\, \Gamma\left(\frac{n-2}{2}\right)}{2^{\frac{n-1}{2}}\, \Gamma\left(\frac{n-1}{2}\right)}$$

$$= \sqrt{\frac{n-1}{2}}\, \frac{\Gamma\left(\frac{n-2}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)}\, \frac{\mu}{\sigma}.$$

Example 5.3.5.2 Let $Z_1, \ldots, Z_n$ be i.i.d. $N(0,1)$ r.v.s and let $a_i, b_i \in \mathbb{R}$, $i = 1, \ldots, n$, be such that $\sum_{i=1}^{n} a_i^2 > 0$, $\sum_{i=1}^{n} b_i^2 > 0$ and $\sum_{i=1}^{n} a_i b_i = 0$. Show that

(a) $$Y_1 = \sqrt{\frac{\sum_{i=1}^{n} b_i^2}{\sum_{i=1}^{n} a_i^2}}\; \frac{\sum_{i=1}^{n} a_i Z_i}{\left|\sum_{i=1}^{n} b_i Z_i\right|} \sim t_1$$

(b) $$Y_2 = \frac{\sum_{i=1}^{n} b_i^2}{\sum_{i=1}^{n} a_i^2} \left(\frac{\sum_{i=1}^{n} a_i Z_i}{\sum_{i=1}^{n} b_i Z_i}\right)^2 \sim F_{1,1}$$

(c) $Y_3 = \sqrt{\dfrac{\sum_{i=1}^{n} b_i^2}{\sum_{i=1}^{n} a_i^2}} \; \dfrac{\sum_{i=1}^{n} a_i z_i}{\sum_{i=1}^{n} b_i z_i} \sim t_1$

<u>Solution</u>

$c_1 \sum_{i=1}^{n} a_i z_i + c_2 \sum_{j=1}^{n} b_j z_j$ a linear combination of $z_1 \ldots z_n$
$\longrightarrow$ univariate normal distribution

$\Rightarrow \left(\sum_{i=1}^{n} a_i z_i , \sum_{j=1}^{n} b_j z_j \right) \sim N_2$

$E\left(\sum_{i=1}^{n} a_i z_i \right) = 0$; $Var\left(\sum_{i=1}^{n} a_i z_i \right) = \sum_{i=1}^{n} a_i^2$

$E\left(\sum_{i=1}^{n} b_i z_i \right) = 0$; $Var\left(\sum_{i=1}^{n} b_i z_i \right) = \sum_{i=1}^{n} b_i^2$

$Cov\left(\sum_{i=1}^{n} a_i z_i , \sum_{j=1}^{n} b_j z_j \right) = \sum_{i=1}^{n} a_i b_i = 0$

$\Rightarrow \left(\sum_{i=1}^{n} a_i z_i , \sum_{j=1}^{n} b_j z_j \right) \sim N_2\left(0, 0, \sum_{i=1}^{n} a_i^2, \sum_{i=1}^{n} b_i^2, 0\right)$

$\Rightarrow \sum_{i=1}^{n} a_i z_i \sim N\left(0, \sum_{i=1}^{n} a_i^2\right)$ and $\sum_{i=1}^{n} b_i z_i \sim N\left(0, \sum_{i=1}^{n} b_i^2\right)$ are independent

$\Rightarrow \dfrac{\sum_{i=1}^{n} a_i z_i}{\sqrt{\sum a_i^2}} \sim N(0,1)$ and $\dfrac{\sum_{i=1}^{n} b_i z_i}{\sqrt{\sum b_i^2}} \sim N(0,1)$ are independent

(a) $\dfrac{\sum_{i=1}^{n} a_i z_i}{\sqrt{\sum_{i=1}^{n} a_i^2}} \sim N\left(0, \sum_{i=1}^{n} a_i^2\right)$ & $\dfrac{\left(\sum_{i=1}^{n} b_i z_i\right)^2}{\sum_{i=1}^{n} b_i^2} \sim \chi_1^2$ are indep't.

$$\Rightarrow \frac{\sum_{i=1}^{n} a_i z_i \Big/ \sqrt{\sum_{i=1}^{n} a_i^2}}{\sqrt{\dfrac{\left(\sum_{i=1}^{n} b_i z_i\right)^2}{\sum_{i=1}^{n} b_i^2}}} \sim t_1$$

i.e. $Y_1 \sim t_1$

(b) Since $t_1^2 \stackrel{d}{=} F_{1,1}$, the result follows on using (a)

(c) $F_{Y_3}(y) = P(Y_3 \le y)$

$= P\left(\frac{Z_1}{Z_2} \le y\right)$, $y \in \mathbb{R}$, (since $Y_3 \stackrel{d}{=} t_1 \stackrel{d}{=} \frac{Z_1}{\sqrt{Y}}$)

where $Z \sim N(0,1)$ and are independent.

Clearly

$F_{Y_3}(y) = P\left(\frac{Z_1}{|Z_2|} \le y, Z_2 > 0\right) + P\left(-\frac{Z_1}{|Z_2|} \le y, Z_2 < 0\right)$

$= P\left(\frac{Z_1}{|Z_2|} \le y, Z_2 > 0\right) + P\left(\frac{Z_1}{|Z_2|} \le y, Z_2 < 0\right)$

$\left((Z_1, Z_2) \stackrel{d}{=} (-Z_1, Z_2)\right)$

$= P\left(\frac{Z_1}{|Z_2|} \le y\right)$, $\forall y \in \mathbb{R}$

$\Rightarrow Y_3 \stackrel{d}{=} \frac{Z_1}{|Z_2|} \sim t_1$ (by (a))

$\Rightarrow Y_3 \sim t_1$.

5.4. Special Multivariate Distributions

5.4.1. Multinomial Distribution
(A generalization of binomial distribution)

$\mathcal{E}$: a random experiment whose each trial results in one (and only one) of $p+1$ possible outcomes $E_1, \ldots, E_{p+1}$ where $E_i \cap E_j = \phi$ and $\sum_{i=1}^{p+1} E_i = \Omega$. Let

$P(A_i) = \theta_i \in (0,1)$, $i = 1, \ldots, k$, and $\sum_{i=1}^{k} \theta_i < 1$ so that

$$P(E_{p+1}) = 1 - \sum_{i=1}^{p} \theta_i \in (0,1)$$

Consider n independent trials of $\mathcal{E}$.

Define

X_i = # of times E_i occurs in n trials, $i = 1, \ldots, p+1$.

Then $\sum_{i=1}^{p+1} X_i = n$, i.e. $X_{p+1} = n - \sum_{i=1}^{n} X_i$. One may

be interested in probability distribution of $\underset{\sim}{X} = (X_1, \dots, X_p)$.
We have

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

$$f_{\underset{\sim}{X}}(x_1, \dots, x_p) = P(X_1 = x_1, \dots, X_p = x_p)$$

$$= \begin{cases} \dfrac{\lfloor n}{\lfloor x_1 \cdots \lfloor x_p \ \lfloor (n - \sum_{i=1}^{p} x_i)} \theta_1^{x_1} \cdots \theta_p^{x_p} \left(1 - \sum_{i=1}^{p} \theta_i\right)^{n - \sum_{i=1}^{p} x_i}, & \underset{\sim}{x} \in S_{\underset{\sim}{X}} \\ 0, & \text{otherwise} \end{cases}$$

$\longrightarrow$ Multinomial distribution with n trials and cell probabilities $\theta_1 \dots \theta_p$ [denoted by $\mathrm{Mult}(n, \theta_1, \dots, \theta_p)$].

$\longrightarrow$ a family of distributions with varying $n \in \mathbb{N}$ and $\underset{\sim}{\theta} = (\theta_1, \dots, \theta_p) \in \Omega = \{(t_1, \dots, t_p) : 0 < t_i < 1\ i = 1, \dots, p \text{ and } \sum_{i=1}^{p} t_i < 1\}$.

Remark 5.4.1.1 for $p = 1$ $\mathrm{Mult}(n, \theta_1)$ distribution is the same as $\mathrm{Bin}(n, \theta_1)$ distribution.

Theorem 5.4.1.1 Suppose that $\underset{\sim}{X} = (X_1, \dots, X_p) \sim \mathrm{Mult}(n, \theta_1, \dots, \theta_p)$, where $n \in \mathbb{N}$ and $\underset{\sim}{\theta} = (\theta_1, \dots, \theta_p) \in \Omega$. Then

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

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

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

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

Proof (a) Fix $i \in \{1, \dots, p\}$. a given trial of the experiment treat the occurrence of E_i as success and its non-occurrence (i.e. 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

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

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

$$X_i + X_j = \#\text{ of successes in } n \text{ Bernoulli trials}$$
$$\sim Bin(n, \theta_i + \theta_j).$$

(c) Obvious

(d) $Var(X_i + X_j) = n(\theta_i+\theta_j)(1-\theta_i-\theta_j)$

$$Var(X_i + X_j) = n(\theta_i+\theta_j)(1-\theta_i-\theta_j)$$
$$\Rightarrow Var(X_i) + Var(X_j) + 2Cov(X_i, X_j) = n(\theta_i+\theta_j)(1-\theta_i-\theta_j)$$
$$\Rightarrow n\theta_i(1-\theta_i) + n\theta_j(1-\theta_j) + 2Cov(X_i, X_j) = n(\theta_i+\theta_j)(1-\theta_i-\theta_j)$$
$$\Rightarrow Cov(X_i, X_j) = -n\theta_i\theta_j$$

MGF The m.g.f. of $\underset{\sim}{X} = (X_1,\dots,X_p)$ is given by

$$M_{\underline{X}}(t_1,\dots,t_p) = E\left(e^{t_1X_1+\dots+t_pX_p}\right)$$
$$= \sum_{\substack{x_1\geq 0 \\ x_1+\dots+x_p\leq n}}^{n} \cdots \sum_{x_p\geq 0}^{n} e^{t_1x_1+\dots+t_px_p} \frac{\lfloor n}{\lfloor x_1 \cdots \lfloor x_p \lfloor n-\sum_{i=1}^p x_i} \theta_1^{x_1}\cdots\theta_p^{x_p}\left(1-\sum_{i=1}^p\theta_i\right)^{n-\sum_{i=1}^p x_i}$$
$$= \sum_{\substack{x_1\geq 0 \\ x_1+\dots+x_p\leq n}}^{n} \cdots \sum_{x_p\geq 0}^{n} \frac{\lfloor n}{\lfloor x_1 \cdots \lfloor x_p \lfloor n-\sum_{i=1}^p x_i} (\theta_1e^{t_1})^{x_1}\cdots(\theta_pe^{t_p})^{x_p}\left(1-\sum_{i=1}^p\theta_i\right)^{n-\sum_{i=1}^p x_i}$$
$$= \left(\theta_1e^{t_1}+\dots+\theta_pe^{t_p}+1-\sum_{i=1}^p\theta_i\right)^n, \quad \underset{\sim}{t} \in \mathbb{R}^p.$$

Remark 5.4.1.2 The last theorem can also be proved using m.g.f.

For example (for $i, j \in \{1,\dots,p\}$, $i\neq j$)

$$M_{X_i+X_j}(t) = M_{\underline{X}}(\underset{\substack{\downarrow \\ i\text{-th position}}}{0,\dots,0,t},0,\dots,0,\underset{\hookrightarrow j\text{-th position}}{t},0\dots) = \underbrace{\left((\theta_i+\theta_j)e^t+1-\theta_i-\theta_j\right)}_{\text{mgf of } Bin(n,\,\theta_i+\theta_j)}, \quad t\in\mathbb{R},$$

5.4.2. Bivariate Normal Distribution

Definition 5.4.2.1 A bivariate r.v. $\underset{\sim}{X} = (X_1, X_2)$ 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 $\underset{\sim}{X} = (X_1, X_2)$ 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 $f_{X_1, X_2}(x_1, x_2) \geq 0$ $\forall \underset{\sim}{x} \in \mathbb{R}^2$ 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

$$\int_{\mathbb{R}^2} f_{X_1, X_2}(x_1, x_2)\, dx_1\, dx_2 = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}(z_1^2 - 2\rho z_1 z_2 + z_2^2)}\, dz_1\, dz_2$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} e^{-\frac{z_2^2}{2}} \Big\{ \underbrace{\int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}(z_1 - \rho z_2)^2}\, dz_1}_{\sqrt{1-\rho^2}\,\sqrt{2\pi}} \Big\}\, dz_2$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z_2^2}{2}}\, dz_2 = 1$$

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

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

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

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

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

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

$X_2 \sim N(\mu_2, \sigma_2^2)$

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

$\mu_1 = E(X_1)$, $\mu_2 = E(X_2)$, $\sigma_1^2 = Var(X_1)$ and $\sigma_2^2 = Var(X_2)$

MGF $\quad M_{X_1,X_2}(t_1, t_2) = E(e^{t_1X_1 + t_2X_2})$

$$= E\left(E(e^{t_1X_1 + t_2X_2} | X_2)\right)$$

$$= E\left(e^{t_2X_2} E(e^{t_1X_1} | X_2)\right), \quad \underset{\sim}{t} = (t_1, t_2) \in \mathbb{R}^2$$

$E(e^{t_1X_1} | X_2) =$ m.g.f. of conditional distn $X_1 | X_2$ at point t_1

$$= e^{\left\{\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(X_2 - \mu_2)\right\}t_1 + \frac{\sigma_1^2(1-\rho^2)t_1^2}{2}}$$

$$\Rightarrow M_{X_1,X_2}(t_1, t_2) = e^{\left\{\mu_1 - \rho\frac{\sigma_1}{\sigma_2}\mu_2\right\}t_1 + \frac{\sigma_1^2(1-\rho^2)t_1^2}{2}} E\left[e^{t_2X_2} e^{\frac{\rho\sigma_1}{\sigma_2}t_1 X_2}\right]$$

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

$$= e^{\left\{\mu_1 - \rho\frac{\sigma_1}{\sigma_2}\mu_2\right\}t_1 + \frac{\sigma_1^2(1-\rho^2)t_1^2}{2}} e^{\mu_2\left\{t_2 + \rho\frac{\sigma_1}{\sigma_2}t_1\right\} + \frac{\sigma_2^2}{2}\left(t_2 + \frac{\rho\sigma_1}{\sigma_2}t_1\right)^2}$$

$$= e^{\mu_1t_1 + \mu_2t_2 + \frac{\sigma_1^2t_1^2}{2} + \frac{\sigma_2^2t_2^2}{2} + \rho\sigma_1\sigma_2t_1t_2}, \quad \underset{\sim}{t} = (t_1, t_2) \in \mathbb{R}^2$$

Thus we have the following theorem.

Theorem 5.4.2.1 Suppose that $\underset{\sim}{X} = (X_1, X_2) \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, $-\infty < \mu_i < \infty$, $\sigma_i > 0$, $i = 1, 2$ 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 $x_2 \in \mathbb{R}$

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

and for fixed $x_1 \in \mathbb{R}$

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

(c) the mgf of $\underset{\sim}{X} = (X_1, X_2)$ is

$$M_{X_1, X_2}(t_1, t_2) = e^{\mu_1 t_1 + \mu_2 t_2 + \frac{\sigma_1^2 t_1^2}{2} + \frac{\sigma_2^2 t_2^2}{2} + \rho\sigma_1\sigma_2 t_1 t_2},$$

$\underset{\sim}{t} = (t_1, t_2) \in \mathbb{R}^2$

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

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

(f) for real constants c_1 and c_2 such that $(c_1, c_2) \neq (0, 0)$

$$c_1 X_1 + c_2 X_2 \sim N(c_1\mu_1 + c_2\mu_2, c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + 2\rho c_1 c_2 \sigma_1 \sigma_2)$$

Solution (a)-(c) Already done

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

$$\psi_{X_1, X_2}(t_1, t_2) = \ln M_{X_1, X_2}(t_1, t_2)$$

$$= \mu_1 t_1 + \mu_2 t_2 + \frac{\sigma_1^2 t_1^2}{2} + \frac{\sigma_2^2 t_2^2}{2} + \rho\sigma_1\sigma_2 t_1 t_2$$

$$\frac{\partial}{\partial t_1}\psi_{X_1, X_2}(t_1, t_2) = \mu_1 + 2\sigma_1^2 t_1 + \rho\sigma_1\sigma_2 t_2$$

$$\frac{\partial^2}{\partial t_2 \partial t_1}\psi_{X_1, X_2}(t_1, t_2) = \rho\sigma_1\sigma_2$$

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

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

(e) Obviously if X_1 and X_2 are independent then

$\rho = Corr(X_1, X_2) = 0$.

Now suppose that $\rho = 0$. Then

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

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

$$= f_{X_1}(x_1)\, f_{X_2}(x_2), \quad \forall\ \underline{x} = (x_1, x_2) \in \mathbb{R}^2$$

$\Rightarrow$ X_1 and X_2 are independent.

(f) Let $Y = c_1X_1 + c_2X_2$. Then

$$M_Y(t) = E(e^{tY}) = E(e^{t(c_1X_1 + c_2X_2)})$$

$$= E(e^{tc_1X_1 + tc_2X_2})$$

$$= M_{X_1, X_2}(tc_1, tc_2)$$

$$= \exp\left\{\mu_1tc_1 + \mu_2tc_2 + \frac{\sigma_1^2t^2c_1^2}{2} + \frac{\sigma_2^2t^2c_2^2}{2} + \rho t^2c_1c_2\sigma_1\sigma_2\right\}$$

$$= \exp\left\{(c_1\mu_1 + c_2\mu_2)t + (c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + 2\rho\sigma_1\sigma_2c_1c_2)\frac{t^2}{2}\right\}$$

$\Longrightarrow$ m.g.f. of $N(c_1\mu_1 + c_2\mu_2,\ c_1^2\sigma_1^2 + c_2^2\sigma_2^2 + 2\rho c_1c_2\sigma_1\sigma_2)$

Theorem 5.4.2.2 Let $\underline{X} = (X_1, X_2)$ be a bivariate r.v. with $E(X_i) = \mu_i \in (-\infty, \infty)$, $Var(X_i) = \sigma_i^2$ $(\sigma_i > 0)$, $i = 1, 2$ and $Corr(X_1, X_2) = \rho \in (-1, 1)$. Then $\underline{X} \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ iff for any $\underline{t} = (t_1, t_2) \in \mathbb{R}^2 - \{0\}$ $Y = t_1X_1 + t_2X_2 \sim N(t_1\mu_1 + t_2\mu_2,\ t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2\rho t_1t_2\sigma_1\sigma_2)$

<u>Proof</u> 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 $\underset{\sim}{t} = (t_1, t_2) \in \mathbb{R}^2 - \{0\}$

$$Y = t_1X_1 + t_2X_2 \sim N(t_1\mu_1 + t_2\mu_2,\ t_1^2\sigma_1^2 + t_2^2\sigma_2^2 + 2\rho t_1t_2\sigma_1\sigma_2)$$

Then, for $\underset{\sim}{t} = (t_1, t_2) \in \mathbb{R}^2 - \{\underset{\sim}{0}\}$

$$M_{X_1,X_2}(t_1, t_2) = E(e^{t_1X_1 + t_2X_2})$$

$$= M_Y(1)$$

$$= e^{t_1\mu_1 + t_2\mu_2 + \frac{t_1^2\sigma_1^2}{2} + \frac{t_2^2\sigma_2^2}{2} + \rho t_1t_2\sigma_1\sigma_2}$$

$\rightarrow$ m.g.f. of $N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$

$$\Rightarrow \underset{\sim}{X} = (X_1, X_2) \sim N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho).$$