Objective :

1. Build a model
2. Estimate the model parameters
3. Evaluate whether one method of estimation is better than the other.
4. Testing of hypothesis.
5. Uncertainty characterization / Interval estimation.

Chapters 5–10, Casella & Berger.

Random sample :

$X_1, \ldots, X_n$ ~~are~~ together is known to be a random sample if they are identically distributed and mutually independent.

Suppose $X_1, \ldots, X_n$ ~~are~~ is a random sample and each X_i has a density $f(x|\theta)$, we mathematically write

$$X_1, \ldots, X_n \overset{iid}{\sim} f(x|\theta).$$

If $X_1, \ldots, X_n \overset{iid}{\sim} f(x|\theta)$, the joint density of $X_1, \ldots, X_n$, given by

$$f(x_1, \ldots, x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$$

Ex: $x_1, \dots, x_n \overset{iid}{\sim} Exp(\lambda)$

then,

$$P(X_1 \le a, x_2 \le a, \dots, x_n \le a)$$

$$= \underbrace{P(X_1 \le a)}\, \underbrace{P(X_2 \le a)} \cdots \underbrace{P(x_n \le a)}$$

$$= [1 - e^{-\lambda a}][1 - e^{-\lambda a}] \cdots [1 - e^{-\lambda a}]$$

$$= [1 - e^{-\lambda a}]^n$$

Result: If $x_1, \dots, x_n$ is a random sample, then $g(x_1), \dots, g(x_n)$ is also a random sample, for any function g.

Results: $x_1, \dots, x_n \overset{iid}{\sim} Pois(\lambda)$, what is the $P(x_1 + \dots + x_n = a)$, a is a positive integer.

Lets say $n = 2$,

$$P(X_1 + X_2 = a) = P\left(\underbrace{\bigcup_{l=0}^{a} \{X_1 = l, X_2 = a - l\}}\right)$$

$$= \sum_{l=0}^{a} P(X_1 = l, X_2 = a - l)$$

$$= \sum_{l=0}^{a} \underbrace{P(X_1 = l)}\, P(X_2 = a - l)$$

$$= \sum_{l=0}^{a} \left\{\frac{e^{-\lambda} \lambda^l}{l!}\right\} \left\{\frac{e^{-\lambda} \lambda^{a-l}}{(a-l)!}\right\}$$

$$= e^{-2\lambda} \lambda^a \sum_{l=0}^{a} \frac{1}{l!\,(a-l)!}$$

$$= \frac{e^{-2\lambda}\lambda^a}{a!} \sum_{l=0}^{a} \frac{a!}{l!\,(a-l)!}$$

$$= \frac{e^{-2\lambda}\lambda^a}{a!} \sum_{l=0}^{a} \binom{a}{l}$$

$$= \frac{e^{-2\lambda}\lambda^a}{a!} 2^a = \frac{e^{-2\lambda}(2\lambda)^a}{a!}$$

$X_1 + X_2 \sim Pois(2\lambda)$

$X_1, X_2 \overset{iid}{\sim} Pois(\lambda) \Rightarrow X_1 + X_2 \sim Pois(2\lambda)$

$X_1 + \cdots + X_n \sim Pois(n\lambda)$

$$P(X_1 + \cdots + X_n = a) = \frac{e^{-n\lambda}(n\lambda)^a}{a!}$$

Some important definitions:

$E[X^k] = \int x^k f(x|\theta)\,dx$, for any k.

$E[X_i X_j] = \iint x_i x_j f(x_i|\theta) f(x_j|\theta)\,dx_i\,dx_j$

$Cov(X_i, X_j) = E[X_i X_j] - E[X_i]\,E[X_j]$
$\quad = 0$ when $X_1, \ldots, X_n$ is a random sample.

The reverse is true only for normal.

Generally, the reverse statement is not true.

Moment Generating Function:

Goal: Calculate $E[X^k]$ for all k integer.

Moment generating function for a random variable X is defined as

$$E\left[e^{tX}\right] = \int e^{tx} f(x|\theta)\, dx = M_X(t)$$

where $f(x|\theta)$ is the density of X.

This is a function of t and it is generally denoted by $M_X(t)$.

Suppose $M_X(t)$ exists in the nbd. of 0.

$$\frac{d}{dt} M_X(t) = \frac{d}{dt} E\left[e^{tX}\right] = E\left[\frac{d}{dt} e^{tX}\right]$$

$$= E\left[X e^{tX}\right]$$

$$\left.\frac{d}{dt} M_X(t)\right|_{t=0} = E[X]$$

$$\frac{d^k}{dt^k} M_X(t) = E\left[X^k e^{tX}\right]$$

$$\Rightarrow \left.\frac{d^k}{dt^k} M_X(t)\right|_{t=0} = E\left[X^k\right]$$

Example: $X \sim N(0,1)$ $\quad E[X] = 0$, $E\left[X^{2k+1}\right] = 0$ for any k integer.

$$E\left[e^{tX}\right] = \int_{-\infty}^{\infty} e^{tx} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\left[x^2 - 2tx\right]}\, dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left[x^2 - 2tx + t^2 - t^2\right]\right\} dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x-t)^2\right\} \exp\left\{\frac{t^2}{2}\right\} dx$$

$$= \exp\left\{\frac{t^2}{2}\right\} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x-t)^2\right\} dx}$$

$$= \exp\left\{\frac{t^2}{2}\right\}$$

$$\frac{d}{dt} \exp\left\{\frac{t^2}{2}\right\}\Bigg|_{t=0} = t \exp\left\{\frac{t^2}{2}\right\}\Bigg|_{t=0} = 0$$

In general check : $X \sim N(\mu, \sigma^2)$

$$M_X(t) = E\left[e^{tX}\right] = \exp\left\{t\mu + \frac{1}{2}t^2\sigma^2\right\}$$

<u>Result</u>:

$X \sim \text{Gamma}(\alpha, \beta)$

$$f(x) = \cancel{\beta^\alpha}\ \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, \quad x > 0$$

$$E\left[e^{tX}\right] = M_X(t) = \int_0^\infty \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} e^{tx}\, dx$$

$$= \int_0^\infty \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} \exp\left\{-x\left(\frac{1}{\beta} - t\right)\right\} dx$$

$$= \int_0^\infty \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} \exp\left\{-\frac{x}{\left\{\frac{1}{\frac{1}{\beta}-t}\right\}}\right\} dx$$

This integral is finite if $\frac{1}{\beta} - t > 0$

$$\Rightarrow t < \frac{1}{\beta}$$

If $t < \frac{1}{\beta}$

$$= \frac{1}{\beta^\alpha \left(\frac{1}{\beta}-t\right)^\alpha} \underbrace{\int_0^\infty \frac{\left(\frac{1}{\beta}-t\right)^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\left\{-\frac{x}{\left\{\frac{1}{\frac{1}{\beta}-t}\right\}}\right\} dx}_{=1}$$

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

Change of variable theorem

$X_1, \ldots, X_n \overset{iid}{\sim} f(x|\theta)$

- Find the joint density of
$U_1 = \Psi_1(X_1, \ldots, X_n)$, $U_2 = \Psi_2(X_1, \ldots, X_n)$, $\ldots$, $U_n = \Psi_n(X_1, \ldots, X_n)$ such that
$(X_1, \ldots, X_n) \longrightarrow (U_1, \ldots, U_n)$ this mapping is one-to-one.

- If $f_U(u_1, \ldots, u_n)$ is the joint density of $U_1, \ldots, U_n$ then

$$f_U(u_1, \ldots, u_n) = \left[\prod_{i=1}^{n} f(H_i(u_1, \ldots, u_n)|\theta)\right] \left|\det\left(\left(\frac{\partial H_i(u_1, \ldots, u_n)}{\partial u_j}\right)_{i,j=1}^{n}\right)\right|$$

Where,

$$X_i = H_i(U_1, \ldots, U_n)$$

<u>Example</u>: $U_1, U_2 \overset{iid}{\sim} U(0,1)$

- $X_1 = \sqrt{-2\log(U_1)}\cos(2\pi U_2)$,

$X_2 = \sqrt{-2\log(U_1)}\sin(2\pi U_2)$.

$$\psi_1(u_1, u_2) = \sqrt{-2\log(u_1)}\,\cos(2\pi u_2)$$

$$\psi_2(u_1, u_2) = \sqrt{-2\log(u_1)}\,\sin(2\pi u_2)$$

$$u_i = H_i(x_1, x_2)$$

$$x_1 = \sqrt{-2\log(u_1)}\,\cos(2\pi u_2)$$

$$x_2 = \sqrt{-2\log(u_1)}\,\sin(2\pi u_2)$$

$$x_1^2 + x_2^2 = -2\log u_1 \quad \Rightarrow \quad u_1 = \exp\left\{-\frac{x_1^2 + x_2^2}{2}\right\}$$

Also,

$$\tan(2\pi u_2) = \frac{x_2}{x_1}$$

$$\Rightarrow \quad u_2 = \frac{1}{2\pi}\tan^{-1}\left(\frac{x_2}{x_1}\right)$$

$$H_1(x_1, x_2) = \exp\left\{-\frac{x_1^2 + x_2^2}{2}\right\}$$

$$H_2(x_1, x_2) = \frac{1}{2\pi}\tan^{-1}\left(\frac{x_2}{x_1}\right)$$

$$\begin{pmatrix} \dfrac{\partial H_1(x_1, x_2)}{\partial x_1} & \dfrac{\partial H_1(x_1, x_2)}{\partial x_2} \\[2ex] \dfrac{\partial H_2(x_1, x_2)}{\partial x_1} & \dfrac{\partial H_2(x_1, x_2)}{\partial x_2} \end{pmatrix}$$

$$J = \begin{pmatrix} -x_1 \exp\left\{-\dfrac{x_1^2 + x_2^2}{2}\right\} & -x_2 \exp\left\{-\dfrac{x_1^2 + x_2^2}{2}\right\} \\[3ex] \dfrac{1}{2\pi}\dfrac{1}{\left(1 + \dfrac{x_2^2}{x_1^2}\right)}\left(-\dfrac{x_2}{x_1^2}\right) & \dfrac{1}{2\pi}\dfrac{1}{\left(1 + \dfrac{x_2^2}{x_1^2}\right)}\left(\dfrac{1}{x_1}\right) \end{pmatrix}$$

$$\det(J) = \frac{1}{2\pi} \frac{1}{\left(1+\frac{x_2^2}{x_1^2}\right)} \exp\left\{-\frac{x_1^2+x_2^2}{2}\right\} (-x_1) \cdot \frac{1}{x_1}$$

$$- \frac{1}{2\pi} \frac{1}{\left(1+\frac{x_2^2}{x_1^2}\right)} \exp\left\{-\frac{x_1^2+x_2^2}{2}\right\} \frac{x_2^2}{x_1^2}$$

$$= -\frac{1}{2\pi} \cancel{\frac{1}{\left(1+\frac{x_2^2}{x_1^2}\right)}} \exp\left\{-\frac{x_1^2+x_2^2}{2}\right\} \cancel{\left(1+\frac{x_2^2}{x_1^2}\right)}$$

$$= -\frac{1}{2\pi} \exp\left\{-\frac{x_1^2+x_2^2}{2}\right\}$$

$$|\det(J)| = \frac{1}{2\pi} \exp\left\{-\frac{x_1^2+x_2^2}{2}\right\}$$

$$f(u_1, u_2) = 1 \quad \text{if} \quad 0 < u_1 < 1, \quad 0 < u_2 < 1$$

$$f_{\underline{X}}(x_1, x_2) = \frac{1}{2\pi} \exp\left\{-\frac{x_1^2+x_2^2}{2}\right\}, \quad \begin{array}{l} -\infty < x_1 < \infty \\ -\infty < x_2 < \infty \end{array}$$

$$f(x_1) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x_1^2}{2}\right\}$$

$$f(x_2) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x_2^2}{2}\right\}$$

$$x_1, x_2 \overset{iid}{\sim} N(0,1)$$

Some important results on random sample

① $X_1, \ldots, X_n$ is a random sample with X_1 having the moment generating function (MGF) $M(t)$. What is the ~~the joint distribution~~ ~~$(X_1, \ldots, X_n)$~~. MGF of

$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$?

check: MGF is $[M(t/n)]^n$

② X and Y are independent random variables with pdf $f_X(x)$ and $f_Y(y)$ respectively. Then the p.d.f of $Z = X+Y$ is
$$f_Z(z) = \int f_X(w) f_Y(z-w)\, dw$$

<u>Pf:-</u> $P(Z \le z) = P(X+Y \le z)$

$$= \int_{-\infty}^{\infty} P(w+Y \le z) f_X(w)\, dw = \int_{-\infty}^{\infty} P(Y \le z-w) f_X(w) dw$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z-w} f_Y(y) f_X(w)\, dy\, dw$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{z} f_Y(y-w) f_X(w) dy\, dw$$

take derivative w.r.t. z on both sides.

③ $X_1, \ldots, X_n \overset{iid}{\sim} N(0, \sigma^2)$

$$\frac{\sum x_i^2}{\sigma^2} \sim \chi^2_n$$

④ $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, let $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$,

$S^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2$. Then

ⓐ $\bar{X}$ and S^2 are independent

ⓑ $\bar{X} \sim N(\mu, \sigma^2/n)$

③ $\frac{(n-1) S^2}{\sigma^2} \sim \chi^2_{n-1}$.

<u>pf:-</u> Define an orthogonal matrix A s.t.

$$A = \begin{pmatrix} \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

For A to be an orthogonal matrix $A'A = AA' = I$

$(a_{i1}, \ldots, a_{in}) \cdot \left(\frac{1}{\sqrt{n}}, \ldots, \frac{1}{\sqrt{n}}\right) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} a_{ij} = 0$

$\sum_{j=1}^{n} a_{ij} = 0$ for all $i = 2, \ldots, n$.

Let $\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ be such that

$$\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$z_1^2 + \cdots + z_n^2 = x_1^2 + \cdots + x_n^2 \quad \cdots\cdots \quad ①$$

$$z_1 = \left(\frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}}\right)\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i = \sqrt{n}\,\bar{x} \quad \cdots \quad ②$$

① & ② together imply

$$z_2^2 + \cdots + z_n^2 = x_1^2 + \cdots + x_n^2 - n\bar{x}^2$$

$$= \sum_{i=1}^{n} (x_i - \bar{x})^2 = (n-1)s^2 \quad \cdots \quad ③$$

$\bar{x}$ is a function of z_1

$(n-1)s^2$ is a function of $z_2, \cdots, z_n$.

$\Rightarrow$ $\bar{x}$ & s^2 are independent

$\bar{x} = \frac{z_1}{\sqrt{n}}$ & we know $z_1 \sim N(\mu, \sigma^2)$

$\Rightarrow$ $\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

and $\frac{(n-1)s^2}{\sigma^2} = \frac{z_2^2 + \cdots + z_n^2}{\sigma^2}$

claim: $z_2, \cdots, z_n \overset{iid}{\sim} N(0, \sigma^2)$

Recap:

① Definition of a random sample

② Some properties of a random sample

③ Change of variable theorem

Result: $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

① $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \sim N(\mu, \frac{\sigma^2}{n})$

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

③ $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

Pf:- $\begin{pmatrix} Z_1 \\ \vdots \\ Z_n \end{pmatrix} = A \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$

A is an orthogonal matrix

$$A = \begin{pmatrix} \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

A orthogonal $\Rightarrow$ $A'A = I$

$\sum_{k=1}^{n} a_{ik}^2 = 1 \;\; \forall i$ & $\sum_{k=1}^{n} a_{ik}a_{jk} = 0 \;\; \forall i \neq j$

$$\sum_{k=1}^{n} a_{ik} = 0 \quad \forall\, i = 2, \dots, n$$

$$z_1^2 + \dots + z_n^2 = x_1^2 + \dots + x_n^2 \quad \dots\dots \quad (1)$$

$$z_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i = \sqrt{n}\, \bar{x}$$

$$\cancel{z_1 \sim N(0, \sigma^2)}$$

From (1), $z_2^2 + \dots + z_n^2 = x_1^2 + \dots + x_n^2 - n\bar{x}^2$

$$\Rightarrow z_2^2 + \dots + z_n^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 = (n-1)s^2$$

Since $\bar{x}$ is a function of z_1 & $(n-1)s^2$ is a function of $z_2, \dots, z_n$, to show that ~~they~~ are independent it is enough to show that z_1 is independent of $z_2, \dots, z_n$.

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \qquad z_1 \sim N(\sqrt{n}\mu, \sigma^2).$$

Goal: z_1 is independent of $z_2, \dots, z_n$.

$$z_1 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i$$

$$z_k = \sum_{i=1}^{n} a_{ki} x_i$$

$$\operatorname{cov}(z_1, z_k) = \operatorname{cov}\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i, \sum_{i=1}^{n} a_{ki} x_i\right)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} a_{ki}\, \sigma^2 = 0$$

$$= \sum_{i=1}^{n} \frac{1}{\sqrt{n}} a_{ki} \operatorname{cov}(x_i, x_i)$$

$\Rightarrow z_1$ is independent of any other z_k

$\Rightarrow z_1$ is independent of $(z_2, \ldots, z_n)$

$\Rightarrow \bar{X}$ is independent of $(n-1)s^2$

(3) To show $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$.

$$(n-1)s^2 = \sum_{i=2}^{n} z_i^2$$

$$z_i = \sum_{k=1}^{n} a_{ik} x_k \qquad \forall\ i = 2, \ldots, n$$

$$E[z_i] = \sum_{k=1}^{n} a_{ik} E[x_k] = \sum_{k=1}^{n} a_{ik}\mu = 0$$

and $\operatorname{Var}(z_i) = \operatorname{Var}\left(\sum_{k=1}^{n} a_{ik} x_k\right)$

$$= \sigma^2 \underbrace{\sum_{k=1}^{n} a_{ik}^2}_{1} = \sigma^2$$

z_i is a linear combination of Normal distributions $\Rightarrow z_i$ is also a Normal distribution

$\Rightarrow z_i \sim N(0, \sigma^2)$

Also, $\operatorname{cov}(z_i, z_j) = \operatorname{Cov}\left(\sum_{k=1}^{n} a_{ik} x_k, \sum_{k=1}^{n} a_{jk} x_k\right)$

$$= \sum_{k=1}^{n} a_{ik} a_{jk} \operatorname{Var}(x_k)$$

$$= \sigma^2 \sum_{k=1}^{n} a_{ik} a_{jk} = 0 \quad \text{As } A \text{ is orthogonal}$$

$\Rightarrow z_i$'s are independent for $i = 2, \ldots, n$

and each $Z_i \sim N(0, \sigma^2)$ $\forall i = 2, \ldots, n$

$$\frac{Z_2^2 + \cdots + Z_n^2}{\sigma^2} \sim \chi^2_{n-1}$$

By ①, $(n-1) s^2 = Z_2^2 + \cdots + Z_n^2$

$$\Rightarrow \frac{(n-1) s^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Result: $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

When σ^2 is known Q What is the distribution of $\sqrt{n}\frac{(\bar{X}-\mu)}{\sigma} \sim N(0,1)$ by part a) of the last result.

$X_1, \ldots, X_n$ $\quad H_0 : \mu = 0$ vs. $H_1 : \mu \neq 0$.

X_i = rainfall in the ith week

You are told rainfall follows a Normal distribution

You are told that the std. deviation of the rainfall is 10 mm. based on some prior study.

Qn: Do the data support the hypothesis that the average rainfall is 45 mm. per week?

$$\left| \frac{\sqrt{n}(\bar{X} - 45)}{10} \right| > 1.96$$ then we reject the hypothesis that the average rainfall is 45 mm.

Qn: You are not told that the standard deviation is 10 mm.

$$\frac{\sqrt{n}\,(\bar{X}-45)}{\sqrt{S^2}} \qquad \text{where} \quad S^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$= \frac{\left.\frac{\sqrt{n}\,(\bar{X}-45)}{\sigma}\right\} \sim N(0,1)}{\sqrt{\frac{S^2}{\sigma^2}}}$$

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

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

$$= \frac{U}{\sqrt{\frac{V}{n-1}}} \qquad \text{where } U \sim N(0,1), \quad V \sim \chi^2_{n-1}$$

and U & V are independent because of the result we have proved in the beginning of the class.

(ii) Start with the joint distribution of (U, V) and find the distribution of $\frac{U}{\sqrt{\frac{V}{n-1}}}$ using the change of variable theorem.

$$f(t) = \frac{\Gamma\left(\frac{n-1+1}{2}\right)}{\Gamma\left(\frac{n-1}{2}\right)} \frac{1}{\sqrt{(n-1)\pi}} \left(1+\frac{t^2}{n-1}\right)^{-n/2}, \quad -\infty < t < \infty$$

This density is known as the ~~t-dis~~ student's t-density with degree of freedom $(n-1)$.

A general student's t-density with degree of freedom p, denoted by t_p, has the density

$$f(t) = \frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma(p/2)} \frac{1}{\sqrt{p\pi}} \left(1+\frac{t^2}{p}\right)^{-\left(\frac{p+1}{2}\right)}, \quad -\infty < t < \infty$$

<u>F-distribution:</u>

$U \sim \chi^2_p$ and $V \sim \chi^2_q$ and U, V are independent then $\frac{U/p}{V/q}$ follows a density known as the $F_{p,q}$ density.

<u>Two sample:</u> $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu_1, \sigma_1^2) \longrightarrow$ sample 1

and $Y_1, \ldots, Y_m \overset{iid}{\sim} N(\mu_2, \sigma_2^2) \longrightarrow$ sample 2

Also sample 1 is independent of sample 2, then, $\dfrac{(n-1)\, s_1^2}{\sigma_1^2} \sim \chi^2_{n-1}$ $\qquad s_1^2 = \dfrac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2$

Similarly $\dfrac{(m-1)\, s_2^2}{\sigma_2^2} \sim \chi^2_{m-1}$ $\qquad s_2^2 = \dfrac{1}{(m-1)} \sum_{j=1}^{m} (y_j - \bar{y})^2$

$$\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} = \frac{\dfrac{\chi^2_{n-1}}{n-1}}{\dfrac{\chi^2_{m-1}}{m-1}}$$

Also s_1^2 and s_2^2 are independent as the two samples are independent.

$$\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim F_{n-1,\, m-1}$$

This result is used to test whether $\sigma_1^2 = \sigma_2^2$.

Properties of F distribution

(a) $X \sim F_{p,q}$ then $\dfrac{1}{X} \sim F_{q,p}$

(b) $X \sim t_q$ then $X^2 \sim F_{1,q}$

(c) $X \sim F_{p,q}$ then $\dfrac{\left(\frac{p}{q}\right) X}{1 + \left(\frac{p}{q}\right) X} \sim \text{Beta}\left(\frac{p}{2}, \frac{q}{2}\right)$