Note on F-distribution and F-test:

# Mathematical Framework:

**F distribution:**

\[ f(F; m, n) = \frac{m^{m/2} n^{n/2}}{(mF+n)^{m+n}/2} \cdot \frac{F^{m/2-1}}{\Gamma(m/2) \Gamma(n/2)} \]

\( m, n \in \mathbb{N} \)

\( F \in \mathbb{R}^+ \)

Since the Beta Function is:

\[ B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} \, dt \]

\[ = \frac{\Gamma(x) \Gamma(y)}{\Gamma(x+y)} = \frac{(x-1)! (y-1)!}{(x+y-1)!} \]

\[ \Rightarrow B(x,y) = B(y,x) \]

Then the other form would be:

\[ f(F; m, n) = \frac{m^{m/2} n^{n/2}}{B(m/2, n/2)} \cdot \frac{F^{m/2-1}}{(mF+n)^{m+n}/2} \]

Or in the explicit form:

\[ f(F; m, n) = \frac{m^{m/2} n^{n/2} (m+n/2 -1)!}{(m/2 -1)! (n/2 -1)!} \cdot \frac{F^{m/2-1}}{(mF+n)^{m+n}/2} \]

*This distribution is often called the Fisher F-distribution, after the famous British statistician Sir Ronald Aylmer Fisher (1890-1962), sometimes the Snedecor F-distribution and sometimes Fisher-Snedecor F-distribution.*
\(* m \) and \(* n \) are called "degrees of freedom"

(2) In statistics, the number of degrees of freedom is the number of independent pieces of data being used to make a calculation.

* It is customary to plot the \( f \) distribution (any of the formulae given for \( f(F; m, n) \) here) against \( F \), for different values of 'n' and 'm'.

![Graphs showing different distributions for \( F \) with \( m = 10 \) and \( n = 10 \).]

*** Mathematical Background:

**Definitions:**

* Probability Mass Function (PMF):

  Probability that "discrete" random variable will exactly equal a discrete value.

* Density function (often called "distribution")
Probability Density Function (PDF):
Probability that "Continuous" random variable will fall within an interval.

* Cumulative Density Function (CDF): 
Probability that either a discrete or continuous random variable will take a value less than or equal to a certain value.

* Characteristic Function:
Fourier transform of the probability density function.

* Chi-distribution:
A direct relation exists between Chi and Gaussian distribution of random variables. If \( x \) is a Gaussian random variable and \( y = x^2 \), then \( y \) has a Chi-distribution with one degree of freedom.

**Notation:**

* Probability Density Function:
\[
\mathbb{P}(r) = \sum_{z=-\infty}^{\infty} P(z) \\
\mathbb{P}(r) = \int_{-\infty}^{r} f(t) \, dt
\]
discrete
continuous

These functions are assumed to be properly normalized:
\[
\sum_{r} \mathbb{P}(r) = 1 \\
\int_{-\infty}^{\infty} f(x) \, dx = 1
\]
\* Characteristic Function:

It is denoted either by \( \Phi \) or \( \Phi \):

\[
\Phi(t) = \int_{-\infty}^{+\infty} e^{itx} f(x) \, dx
\]

The Inverse: 
\[f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Phi(t) e^{-itx} \, dt\]

++ In some texts, characteristic Function is denoted as a function of \( \epsilon x \). Some of the notations include it as a function of frequency \( \omega \),

\[\Phi(j\omega) = \int_{-\infty}^{+\infty} f(x) e^{j\omega x} \, dx\]

where \( j \) is used for \( \sqrt{-1} \) instead of \( i \) to avoid confusion (with \( i \) index).

\* Derivation of the Chi-Square distribution:

We work our way out, with the Cumulative Density Function (CDF). CDF is denoted by \( F(x) \):

\[F(x) = P(X \leq x)\]

Obviously: \( 0 \leq F(x) \leq 1 \)

\[F(-\infty) = 0, \quad F(\infty) = 1\]

**EX.**

In case of flipping a coin:

\[X(s) = \begin{cases} 
1 & s = \text{H} \\
-1 & s = \text{T} 
\end{cases}\]
CDF plot:

Probability density function is related to CDF by the following equation:

\[
dF(x) = p(x) \quad -\infty < x < +\infty
\]

\[F(x) = \int_{-\infty}^{x} p(u) du\]

Probability of \( X \) between \( x_1 \) and \( x_2 \) (\( x_2 > x_1 \))

\[
P(x_1 < X \leq x_2) = P(X \leq x_1) + P(x_1 < X < x_2)
\]

\[F(x_2) - F(x_1) = P(x_1 < X < x_2)\]

*To derive the Chi distribution equation, let's start with a simple case and go on from there.*

\( X, Y \) = random variables

1) \( Y = aX + b \quad a, b \) : constants
\[ F_X(x) = \text{CDF for } X \]
\[ F_Y(y) = \text{CDF for } Y \]

\[ F_Y(y) = P(Y < y) = P(aX + b < y) = P\left( X < \frac{y - b}{a} \right) \]

\[ F_Y(y) = \int_{-\infty}^{\frac{y-b}{a}} F_X(x) \, dx = F_X\left( \frac{y-b}{a} \right) \]

\[ \Rightarrow \quad \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\left( \frac{y-b}{a} \right) \]

\[ \Rightarrow \quad p_Y(y) = \frac{1}{a} p_X\left( \frac{y-b}{a} \right) \]

2) We now will calculate for a more complex case:

\[ Y = aX^2 + b \]

\[ F_Y(y) = P(Y < y) = P(aX^2 + b < y) \]

\[ = P(1X^2 < \sqrt{\frac{y-b}{a}}) = \sqrt{\frac{y-b}{a}} \]

\[ = \int_{-\infty}^{-\sqrt{\frac{y-b}{a}}} p_X(u) \, du + \int_{-\infty}^{\infty} p_X(u) \, du \]

\[ - \int_{-\infty}^{-\sqrt{\frac{y-b}{a}}} p_Y(u) \, du \]

\[ = -F_X\left( -\sqrt{\frac{y-b}{a}} \right) + F_X\left( \sqrt{\frac{y-b}{a}} \right) \]

\[ \Rightarrow \quad \frac{d}{dy} F_Y(y) = p_Y(y) = \frac{1}{2\sqrt{\frac{y-b}{a}}} p_X\left( -\sqrt{\frac{y-b}{a}} \right) + \frac{1}{2\sqrt{\frac{y-b}{a}}} p_X\left( \sqrt{\frac{y-b}{a}} \right) \]

Finally, in order to obtain the solution for chi distribution we consider \( Y = X^2 \).
which means: \( a = 1 \), \( b = 0 \)

\[
\Rightarrow P_Y(y) = \frac{1}{2\sqrt{\theta}} P_X(-\sqrt{\theta}) + \frac{1}{2\sqrt{\theta}} P_X(\sqrt{\theta})
\]

Now if we assume the general form of a Gaussian distribution to be:

\[
G(x) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}
\]

by assuming \( a = 0 \) (without losing generality) we obtain:

\[
P_X(\sqrt{\theta}) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{\theta/2\sigma^2}{2\sigma^2}}, \quad P_X(-\sqrt{\theta}) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{-\theta/2\sigma^2}{2\sigma^2}}
\]

Then:

\[
P_Y(y) = \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{y/2\sigma^2}{2\sigma^2}}
\]

Finally the characteristic function would be:

\[
F_Y(\omega) = \int_{-\infty}^{\infty} P_Y(y) dy = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{y^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{-y/2\sigma^2}{2\sigma^2}} dy = \frac{1}{\sqrt{2\pi \sigma^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \sigma^2}} e^{-\frac{-y/2\sigma^2}{2\sigma^2}} dy = \frac{1}{\sqrt{2\pi \sigma^2}}
\]

\[
\Rightarrow \mathcal{F}(\hat{\omega}) = \frac{1}{\sqrt{1 - 2\hat{\omega}^2 \sigma^2}}
\]
The inverse transform of the obtained characteristic function leads to probability density function:

\[ P_Y(y) = \frac{1}{\sqrt{2\pi} \sigma} y^{\frac{\nu}{2} - 1} e^{-\frac{y}{2\sigma^2}} \]

Since for a system with \( \nu \) degrees of freedom we have:

\[ g(z) = \frac{1}{(1-z^2\sigma^2)^{\nu/2}} \]

(obtained in an identical way to the 1 degree of freedom system, just assume: \( Y = \sum_{251}^{\infty} X_i^2 \))

thus:

\[ P_Y(y) = \frac{1}{\sqrt{2\pi} 2^{\nu/2} \Gamma(\nu/2)} y^{\frac{\nu}{2} - 1} e^{-\frac{y}{2\sigma^2}} \]

which is called a chi-square pdf with \( \nu \) degrees of freedom.

---

END of Mathematical Background

A practical example of F-distribution is given in the following
Variance Ratio:

we could use F-distribution when estimates of
the variance for two independent samples from normal
distributions

\[ S_1^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2, \quad S_2^2 = \frac{1}{n-1} \sum_{j=1}^{n} (Y_j - \bar{Y})^2 \]

have been made.

Here \( S_1^2 \) and \( S_2^2 \) are estimates of \( \sigma_1^2 \) and \( \sigma_2^2 \).
That is to say \( (n-1)S_1^2 / \sigma_1^2 \) and \( (n-1)S_2^2 / \sigma_2^2 \) are
distributed according to the chi-square distribution
with \( n-1 \) and \( n-1 \) degrees of freedom respectively.

In this case the quantity

\[ F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} \]

is distributed according to the F-distribution with
\( n-1 \) and \( n-1 \) degrees of freedom.

If the true variances of the two populations are
indeed the same, then the variance ratio \( S_1^2 / S_2^2 \) has
We would reject the null hypotheses at the \( \alpha \) confidence level if the F-ratio is less than \( F_{(1-\alpha/2), (m-1), (n-1)} \) or greater than \( F_{(\alpha/2), (m-1), (n-1)} \) where \( F_{a, m, n} \) is defined by:

\[
\int_{0}^{F_{a, m, n}} f(F; m, n) \, dF = 1 - \alpha
\]

which

where \( \alpha \) would be the probability content of the distribution above the value \( F_{a, m-1, n-1} \).