STAT 24400 Lecture 13
Section 6.2 χ^2, t, and F Distributions
Section 6.3 Sample Mean & Sample Variance

Yibi Huang
Department of Statistics
University of Chicago

Section 6.2 χ^2, t, and F Distributions

Chapter 6 Distributions Derived from Normal

There are 3 distributions derived from from the normal distributions that occur many statistical problems

- Chi-Squared (χ^2) distributions
  - "Chi-squared" is read "kai-squared"
- t distributions
- F distributions

Definitions: Chi-Squared Distributions

Let $Z_1, Z_2, \dots, Z_n$ be i.i.d. $\sim N(0,1)$. The random variable

$$T_n = \sum_{i=1}^{n} Z_i^2$$

is said to be a ***chi-squared distribution*** with n *degrees of freedom*, denoted as

$$T_n \sim \chi_n^2.$$

In HW9, we show using MGF that chi-squared distributions are special Gamma distributions that

$$\chi_n^2 = \text{Gamma}(\alpha = n/2, \lambda = 1/2)$$

and the corresponding PDF is

$$f_{T_n}(t) = \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} t^{(n/2)-1} e^{-t/2}, \quad t > 0.$$

Properties of Chi-Squared Distributions

If $Y \sim \chi^2_n$, then its MGF is

$$M(t) = (1 - 2t)^{-n/2},$$

from which we can derive its expected value and variance

- $\mathrm{E}[Y] = n$
- $\mathrm{Var}(Y) = 2n$
- If $U \sim \chi^2_n$ and $V \sim \chi^2_m$ are independent, then $U + V \sim \chi^2_{m+n}$
  - The proof is straight forward using MGF

Definition: (Student's) t-Distributions

If $Z \sim N(0,1)$ and $U \sim \chi_n^2$ and Z and U are independent, then the distribution of

$$T = \frac{Z}{\sqrt{U/n}}$$

is called the ***(Student's) t-distribution*** with n *degrees of freedom*, denoted as

$$T \sim t_n.$$

The PDF is given by

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

Proof for the PDF of the t-Distribution

By the independence of Z and U, their joint PDF is given by

$$f_{ZU}(z,u) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \cdot \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} u^{\frac{n}{2}-1} e^{-u/2} = \frac{u^{\frac{n}{2}-1}\exp(-\frac{1}{2}(z^2+u))}{\sqrt{\pi}2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})}, \quad -\infty < z < \infty,\ u > 0.$$

Consider the transformation $W = \frac{Z}{\sqrt{U}}$, $Y = U$, with inverse transformation

$$\begin{aligned} Z &= W\sqrt{Y}, \\ U &= Y \end{aligned} \quad \Rightarrow \text{Jacobian} = \begin{vmatrix} \frac{\partial z}{\partial w} & \frac{\partial z}{\partial y} \\ \frac{\partial u}{\partial w} & \frac{\partial u}{\partial y} \end{vmatrix} = \begin{vmatrix} \sqrt{y} & \frac{w}{2\sqrt{y}} \\ 0 & 1 \end{vmatrix} = \sqrt{y}.$$

The joint PDF for (W,Y) is

$$\begin{aligned} f_{WY}(w,y) &= f_{ZU}(w\sqrt{y}, y) \cdot \sqrt{y} \\ &= \frac{y^{\frac{n}{2}-1}\exp(-\frac{1}{2}(w^2y+y))}{\sqrt{\pi}2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})}\sqrt{y} = \frac{y^{\frac{n+1}{2}-1}\exp\left(-\frac{y}{2}\left(1+w^2\right)\right)}{\sqrt{\pi}2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})} \end{aligned}$$

The marginal PDF for W can be obtained by integrating $f_{WY}(w, y)$ over y.

$$f_W(w) = \int_0^\infty f_{WY}(w, y)\mathrm{d}y = \frac{1}{\sqrt{\pi}2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})}\int_0^\infty y^{\frac{n+1}{2}-1}e^{-\frac{y}{2}(1+w^2)}\mathrm{d}y.$$

Let

$$x = \frac{y}{2}(1 + w^2) \ \Rightarrow y = \frac{2x}{1 + w^2},\ \mathrm{d}y = \frac{2}{1 + w^2}\mathrm{d}x.$$

Then,

$$\begin{aligned} f_W(w) &= \frac{1}{\sqrt{\pi}2^{\frac{n+1}{2}}\Gamma(\frac{n}{2})}\int_0^\infty \left(\frac{2x}{1 + w^2}\right)^{\frac{n+1}{2}-1} e^{-x}\frac{2}{1 + w^2}\mathrm{d}x \\ &= \frac{1}{\sqrt{\pi}\Gamma(\frac{n}{2})(1 + w^2)^{\frac{n+1}{2}}} \underbrace{\int_0^\infty x^{\frac{n+1}{2}-1}e^{-x}\mathrm{d}x}_{=\Gamma(\frac{n+1}{2})} \\ &= \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi}\Gamma(\frac{n}{2})}(1 + w^2)^{-\frac{n+1}{2}}, \quad -\infty < w < \infty. \end{aligned}$$

The PDF for $T = \frac{Z}{\sqrt{U/n}} = \sqrt{n}\,W$ is

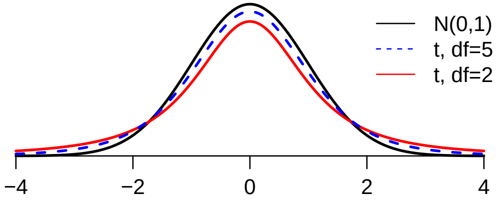
$$f_T(t) = \frac{1}{\sqrt{n}} f_W\left(\frac{t}{\sqrt{n}}\right) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}\left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}, \quad -\infty < t < \infty.$$

Properties of t-Distributions

For $T \sim t_n$ with the PDF

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

- Bell-shaped, symmetric about 0
- With 1 degree of freedom, t_1 = Cauchy
- $\mathrm{E}[T] = 0$ if df > 1
- For large t, the t-density with n df is $\approx \frac{\text{constant}}{t^{n+1}}$ $\Rightarrow$ heavier tail than normal
- $\mathrm{E}[T^k]$ doesn't exist if $k \geq$ degrees of freedom (df)
- higher df $\Rightarrow$ lighter tails
- As df $\to \infty$, $t \to N(0,1)$



Definition: F-Distributions

Let U and V be independent chi-square random variables with m and n degrees of freedom, respectively. The distribution of

$$X = \frac{U/m}{V/n}$$

is called the *F-distribution with m and n degrees of freedom*, denoted by

$$X \sim F_{m,n}.$$

The PDF is given by

$$f(x) = \frac{\Gamma(\frac{m+n}{2})}{\Gamma(\frac{m}{2})\Gamma(\frac{n}{2})} \left(\frac{m}{n}\right)^{m/2} x^{\frac{m}{2}-1} \left(1 + \frac{m}{n}x\right)^{-\frac{m+n}{2}}, \quad x > 0.$$

The PDF can be obtained similarly as for the t-distribution.

Properties of F-Distributions

For $X \sim F_{m,n}$

- $\mathrm{E}(X) = \dfrac{n}{n-2}$ if $n > 2$
- $\mathrm{E}(X^k)$ exists only if $k < n/2$
- If $T \sim t_n$, then $T^2 \sim F_{1,n}$
- asymmetric PDF
- F-distribution can be transformed to Beta distribution

$$X \sim F_{m,n} \quad \Rightarrow \quad Y = \frac{(m/n)X}{1 + (m/n)X} \sim \text{Beta}\left(a = \frac{n}{2}, b = \frac{m}{2}\right)$$

Section 6.3 Sample Mean & Sample Variance

First Statistics Question in STAT 24400

If we observed

$$X_1, X_2, \dots, X_n, \text{ i.i.d.} \sim N(\mu, \sigma^2), \quad \text{but } \mu \text{ and } \sigma^2 \text{ are UNKNOWN.}$$

How to use the observed values of $X_1, X_2, \dots, X_n$ to estimate the unknown μ and σ^2?

First Statistics Question in STAT 24400

If we observed

$$X_1, X_2, \ldots, X_n, \text{ i.i.d.} \sim N(\mu, \sigma^2), \quad \text{but } \mu \text{ and } \sigma^2 \text{ are UNKNOWN.}$$

How to use the observed values of $X_1, X_2, \ldots, X_n$ to estimate the unknown μ and σ^2?

- $X_1, X_2, \ldots, X_n$ are sometimes called the *sample*, n is called the *sample size*
- Usually estimate μ by $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, called the *sample mean*.
- As $\sigma^2 = \mathrm{E}[(X_i - \mu)^2]$, one might attempt to estimate it by

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

  However, μ is unknown. We thus estimate σ^2 by

$$S^2 = \frac{\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{n-1}, \quad \text{called the } \textit{sample variance}.$$

  - Why divide by $n-1$, not n?
- We will discuss estimation problems in Chapter 8 in detail

Population Mean/Variance v.s. Sample Mean/Variance

If $X_1, \ldots, X_n$ are i.i.d. $\sim N(\mu, \sigma^2)$,

- μ is called the *population mean*
- $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ is called the *sample mean*
- σ^2 is called the *populaition variance*
- $S^2 = \dfrac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1}$ is called the *sample variance*

Sample Mean

For the sample mean

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

we have shown earlier that

$$\mathrm{E}(\overline{X}) = \mu \quad \text{and} \quad \mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}$$

and

$$\overline{X} \sim N\left(\mu,\, \frac{\sigma^2}{n}\right).$$

A Useful Identity

The following identity always holds for any value of c.

$$\sum_{i=1}^{n}(X_i - c)^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2 + n(\overline{X} - c)^2, \quad \text{where } \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Proof.

$$\begin{aligned}
\textstyle\sum_{i=1}^{n}(X_i - c)^2 &= \textstyle\sum_{i=1}^{n}(X_i - \overline{X} + \overline{X} - c)^2 \\
&= \textstyle\sum_{i=1}^{n}(X_i - \overline{X})^2 + 2\sum_{i=1}^{n}(X_i - \overline{X})\underbrace{(\overline{X} - c)}_{\text{constant}} + \sum_{i=1}^{n}\underbrace{(\overline{X} - c)^2}_{\text{constant}} \\
&= \textstyle\sum_{i=1}^{n}(X_i - \overline{X})^2 + 2(\overline{X} - c)\underbrace{\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})}_{=0,\ \text{see below}} + n(\overline{X} - c)^2
\end{aligned}$$

where $\displaystyle\sum_{i=1}^{n}(X_i - \overline{X}) = \sum_{i=1}^{n} X_i - \sum_{i=1}^{n}\overline{X} = n\overline{X} - n\overline{X} = 0.$

Corollary of the Useful Identity

$$\sum_{i=1}^{n}(X_i - c)^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2 + n(\overline{X} - c)^2$$

- The case $c = 0$ gives the shortcut formula for the sample variance

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1} = \frac{\left(\sum_{i=1}^{n} X_i^2\right) - n\overline{X}^2}{n-1}.$$

- The value c that minimizes $\sum_{i=1}^{n}(X_i - c)^2$ is $c = \overline{X}$.

Expectation of Sample Variance (Why Divide by $n-1$, not n?)

Letting $c = \mu = \mathrm{E}[X_i]$ in the useful identity

$$\sum_{i=1}^{n}(X_i-\mu)^2 = \underbrace{\sum_{i=1}^{n}(X_i-\overline{X})^2}_{=(n-1)S^2} + n(\overline{X}-\mu)^2.$$

gives the following expression for S^2

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n}(X_i-\mu)^2 - n(\overline{X}-\mu)^2\right).$$

Taking expected values on both sides, we get

$$\begin{aligned}\mathrm{E}[S^2] &= \frac{1}{n-1}\left(\sum_{i=1}^{n}\underbrace{\mathrm{E}[(X_i-\mu)^2]}_{=\mathrm{Var}(X_i)=\sigma^2} - n\underbrace{\mathrm{E}[(\overline{X}-\mu)^2]}_{=\mathrm{Var}(\overline{X})=\sigma^2/n}\right)\\ &= \frac{1}{n-1}\left(n\sigma^2 - n\cdot\frac{\sigma^2}{n}\right) = \sigma^2.\end{aligned}$$

$\overline{X}$ is Independent of S^2

$\overline{X}$ Is Independent of S^2 ($\star\star$)

We will first prove that

$$\overline{X} \quad \text{is indep. of} \quad \overbrace{(X_{2} - \overline{X}, X_3 - \overline{X}, \dots, X_n - \overline{X})}^{\text{no } X_1 - \overline{X}}.$$

This would imply $\overline{X}$ is independent of S^2 since $(n-1)S^2$ can be written as a function of $(X_2 - \overline{X}, X_3 - \overline{X}, \dots, X_n - \overline{X})$ as follows

$$(n-1)S^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2 = (\underbrace{X_1 - \overline{X}}_{\text{See below}})^2 + \sum_{i=2}^{n}(X_i - \overline{X})^2$$

where $X_1 - \overline{X} = -\sum_{i=2}^{n}(X_i - \overline{X})$ since $\sum_{i=1}^{n}(X_i - \overline{X}) = 0.$

$\overline{X}$ Is Independent of S^2 (★★)

We will first prove that

$$\overline{X} \quad \text{is indep. of} \quad \overbrace{(X_2 - \overline{X}, X_3 - \overline{X}, \ldots, X_n - \overline{X})}^{\text{no } X_1 - \overline{X}}.$$

This would imply $\overline{X}$ is independent of S^2 since $(n-1)S^2$ can be written as a function of $(X_2 - \overline{X}, X_3 - \overline{X}, \ldots, X_n - \overline{X})$ as follows

$$(n-1)S^2 = \sum_{i=1}^{n}(X_i - \overline{X})^2 = (\underbrace{X_1 - \overline{X}}_{\text{See below}})^2 + \sum_{i=2}^{n}(X_i - \overline{X})^2$$

where $X_1 - \overline{X} = -\sum_{i=2}^{n}(X_i - \overline{X})$ since $\sum_{i=1}^{n}(X_i - \overline{X}) = 0$.

Steps of the proof:

1. find the joint PDF $f_{\mathbf{Y}}(y_1, y_2, \ldots, y_n)$ of $Y_1 = \overline{X}$, $Y_i = X_i - \overline{X}$ for $i = 2, \ldots, n$.
2. show that the joint PDF $f_{\mathbf{Y}}(y_1, y_2, \ldots, y_n)$ can factor as the product of a function of y_1 and a function of $(y_2, \ldots, y_n)$.

$$f(y_1, y_2, \ldots, y_n) = g(y_1)h(y_2, \ldots, y_n), \quad \text{for all } y_1, y_2, \ldots, y_n.$$

Multivariate Transformation

Suppose $(X_1, \ldots, X_n)$ are continuous r.v.'s with joint PDF

$$f_{\mathbf{X}}(x_1, \ldots, x_n).$$

They are mapped onto $(Y_1, \ldots, Y_n)$ by a 1-to-1 transformation

$$\begin{aligned} y_1 &= g_1(x_1, \ldots, x_n) \\ &\vdots \\ y_n &= g_n(x_1, \ldots, x_n) \end{aligned}$$

and the transformation can be inverted to obtain

$$\begin{aligned} x_1 &= h_1(y_1, \ldots, y_n) \\ &\vdots \\ x_n &= h_n(y_1, \ldots, y_n). \end{aligned}$$

The joint PDF $f_{\mathbf{Y}}(y_1, \dots, y_n)$ is given by

$$f_{\mathbf{Y}}(y_1, \dots, y_n) = f_{\mathbf{X}}(h_1(y_1, \dots, y_n), \dots, h_n(y_1, \dots, y_n)) \left| \frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)} \right|,$$

where $\left| \frac{\partial(x_1, \dots, x_n)}{\partial(y_1, \dots, y_n)} \right|$ is absolute value of the *Jacobian of the transformation*, defined as the determinant of the $n \times n$ matrix

$$\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \vdots & \ddots & \vdots & \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$$

that the (i, j) element is $\dfrac{\partial x_i}{\partial y_j}$.

Joint PDF of $\overline{X}$ and $(X_2 - \overline{X}, \dots, X_n - \overline{X})$

For $Y_1 = \overline{X}$, $Y_i = X_i - \overline{X}$, for $i = 2, 3, \dots, n$, the inverse transformation is

$$X_1 = Y_1 - (Y_2 + Y_3 + \cdots + Y_n),$$
$$X_i = Y_1 + Y_i, \quad \text{for } i = 2, 3, \dots, n.$$

We see

$$\frac{\partial x_1}{\partial y_j} = \begin{cases} 1 & \text{if } j = 1 \\ -1 & \text{if } j = 2, 3, \dots, n, \end{cases} \quad \text{and} \quad \frac{\partial x_i}{\partial y_j} = \begin{cases} 1 & \text{if } j = 1 \text{ or } i \\ 0 & \text{otherwise.} \end{cases}$$

The Jacobian matrix is

$$\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \frac{\partial x_1}{\partial y_3} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \frac{\partial x_2}{\partial y_3} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \frac{\partial x_3}{\partial y_1} & \frac{\partial x_3}{\partial y_2} & \frac{\partial x_3}{\partial y_3} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \frac{\partial x_n}{\partial y_3} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix} = \begin{vmatrix} 1 & -1 & -1 & \cdots & -1 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{vmatrix}$$

The determinant can be shown by induction to be n.

As X_i's are independent, their joint PDF is

$$f_{\mathbf{X}}(x_1, x_2, \dots, x_n) = \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left(\frac{-1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2\right)$$
$$= \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left(\frac{-1}{2\sigma^2} \left(\sum_{i=1}^{n} (x_i - \overline{x})^2 + n(\overline{x} - \mu)^2\right)\right)$$

in which, $$x_1 - \overline{x} = -\sum_{i=2}^{n} (x_i - \overline{x}) = -\sum_{i=2}^{n} y_i$$

$$\sum_{i=1}^{n} (x_i - \overline{x})^2 = (x_1 - \overline{x})^2 + \sum_{i=2}^{n} (x_i - \overline{x})^2 = \left(\sum_{i=2}^{n} y_i\right)^2 + \sum_{i=2}^{n} y_i^2$$

The joint PDF of $(Y_1, \dots, Y_n)$ is thus

$$f_{\mathbf{Y}}(y_1, y_2, \dots, y_n) = \frac{|J|}{(2\pi)^{n/2}\sigma^n} \exp\left[\frac{-1}{2\sigma^2} \left((\sum_{i=2}^{n} y_i)^2 + \sum_{i=2}^{n} y_i^2 + n(y_1 - \mu)^2 \right)\right]$$

where $|J| = n$ is the Jacobian shown on the previous page.

We can see the joint PDF $f_{\mathbf{Y}}(y_1, y_2, \dots, y_n)$ can factor into

- a function $\exp(-\frac{n}{2\sigma^2}(y_1 - \mu)^2)$ of y_1, and
- a function $\exp\left[\frac{-1}{2\sigma^2}\left((\sum_{i=2}^n y_i)^2 + \sum_{i=2}^n y_i^2\right)\right]$ of $y_2, \dots, y_n$.

This proves the independence of

$$Y_1 = \overline{X} \quad \text{and} \quad (Y_2, \dots, Y_n) = (X_2 - \overline{X}, \dots, X_n - \overline{X}),$$

which implies the independence of $\overline{X}$ and S^2.

Distribution of S^2

If $X_1, X_2, \dots, X_n$ are i.i.d. $\sim N(\mu, \sigma^2)$, then

$$\frac{X_1 - \mu}{\sigma}, \frac{X_2 - \mu}{\sigma}, \dots, \frac{X_n - \mu}{\sigma} \quad \text{are i.i.d.} \quad \sim N(0,1),$$

which implies

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

has a **chi-squared** distribution with n degrees of freedom.

If $X_1, X_2, \dots, X_n$ are i.i.d. $\sim N(\mu, \sigma^2)$, then

$$\frac{X_1 - \mu}{\sigma}, \frac{X_2 - \mu}{\sigma}, \dots, \frac{X_n - \mu}{\sigma} \quad \text{are i.i.d.} \quad \sim N(0, 1),$$

which implies

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

has a **chi-squared** distribution with n degrees of freedom.

Question: What's the distribution of

$$\sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2}?$$

If $X_1, X_2, \dots, X_n$ are i.i.d. $\sim N(\mu, \sigma^2)$, then

$$\frac{X_1 - \mu}{\sigma}, \frac{X_2 - \mu}{\sigma}, \dots, \frac{X_n - \mu}{\sigma} \quad \text{are i.i.d.} \quad \sim N(0, 1),$$

which implies

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

has a **chi-squared** distribution with n degrees of freedom.

Question: What's the distribution of

$$\sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2}?$$

Ans: **chi-squared** distribution with $n - 1$ degrees of freedom.

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

Define V_1, V_2, V_3 as follows:

$$\underbrace{\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}}_{=V_1} = \underbrace{\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma^2}}_{=V_2} + \underbrace{\frac{n(\overline{X}-\mu)^2}{\sigma^2}}_{=V_3}.$$

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

Define V_1, V_2, V_3 as follows:

$$\underbrace{\frac{\sum_{i=1}^n (X_i-\mu)^2}{\sigma^2}}_{=V_1} = \underbrace{\frac{\sum_{i=1}^n (X_i-\overline{X})^2}{\sigma^2}}_{=V_2} + \underbrace{\frac{n(\overline{X}-\mu)^2}{\sigma^2}}_{=V_3}.$$

- From the previous page, $V_1 \sim \chi^2_n$ has MGF $M_{V_1}(t) = (1-2t)^{-n/2}$

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

Define V_1, V_2, V_3 as follows:

$$\underbrace{\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}}_{=V_1} = \underbrace{\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma^2}}_{=V_2} + \underbrace{\frac{n(\overline{X}-\mu)^2}{\sigma^2}}_{=V_3}.$$

- From the previous page, $V_1 \sim \chi^2_n$ has MGF $M_{V_1}(t) = (1-2t)^{-n/2}$
- $\sqrt{n}(\overline{X}-\mu)/\sigma \sim N(0,1) \;\Rightarrow V_3 \sim \chi^2_1$ with MGF $M_{V_3}(t) = (1-2t)^{-1/2}$

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

Define V_1, V_2, V_3 as follows:

$$\underbrace{\frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2}}_{=V_1} = \underbrace{\frac{\sum_{i=1}^n (X_i - \overline{X})^2}{\sigma^2}}_{=V_2} + \underbrace{\frac{n(\overline{X} - \mu)^2}{\sigma^2}}_{=V_3}.$$

- From the previous page, $V_1 \sim \chi^2_n$ has MGF $M_{V_1}(t) = (1-2t)^{-n/2}$
- $\sqrt{n}(\overline{X} - \mu)/\sigma \sim N(0,1) \Rightarrow V_3 \sim \chi^2_1$ with MGF $M_{V_3}(t) = (1-2t)^{-1/2}$
- Indep of V_2 and V_3 comes from the indep of S^2 and $\overline{X}$.
  The MGF of $V_1 = V_2 + V_3$ is thus

  $$M_{V_1}(t) = M_{V_2}(t) M_{V_3}(t)$$

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

Define V_1, V_2, V_3 as follows:

$$\underbrace{\frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2}}_{=V_1} = \underbrace{\frac{\sum_{i=1}^n (X_i - \overline{X})^2}{\sigma^2}}_{=V_2} + \underbrace{\frac{n(\overline{X} - \mu)^2}{\sigma^2}}_{=V_3}.$$

- From the previous page, $V_1 \sim \chi^2_n$ has MGF $M_{V_1}(t) = (1-2t)^{-n/2}$
- $\sqrt{n}(\overline{X} - \mu)/\sigma \sim N(0,1) \ \Rightarrow V_3 \sim \chi^2_1$ with MGF $M_{V_3}(t) = (1-2t)^{-1/2}$
- Indep of V_2 and V_3 comes from the indep of S^2 and $\overline{X}$.
  The MGF of $V_1 = V_2 + V_3$ is thus

  $$M_{V_1}(t) = M_{V_2}(t) M_{V_3}(t) \ \Rightarrow \ M_{V_2}(t) = \frac{M_{V_1}(t)}{M_{V_3}(t)} = \frac{(1-2t)^{-n/2}}{(1-2t)^{-1/2}} = (1-2t)^{-\frac{n-1}{2}},$$

  which is the MGF for χ^2_{n-1}. By the uniqueness of MGFs, this proves

  $$V_2 = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{\sigma^2} = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

Back to Statistics

Recall our goal is to estimate the unknown mean μ using the observed values of $X_1, X_2, \dots, X_n$ that are i.i.d. $\sim N(\mu, \sigma^2)$.

Back to Statistics

Recall our goal is to estimate the unknown mean μ using the observed values of $X_1, X_2, \ldots, X_n$ that are i.i.d. $\sim N(\mu, \sigma^2)$.

For $Z \sim N(0,1)$, using the normal CDF we know

$$\mathrm{P}(-1.96 \le Z \le 1.96) = 0.95.$$

As $\overline{X} \sim N(\mu, \sigma^2/n)$, which implies $\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$, we have

$$\mathrm{P}\left(-1.96 \le \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \le 1.96\right) = 0.95,$$

or equivalently

$$\mathrm{P}\left(\overline{X} - 1.96\frac{\sigma}{\sqrt{n}} \le \mu \le \overline{X} + 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95.$$

This means, for 95% of the time, the sample mean $\overline{X}$ is within $1.96\sigma/\sqrt{n}$ from the true value of μ, but σ is UNKNOWN.

t-Statistic

The result on the previous page relies on the fact that

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}} \sim N(0,1), \quad \text{but } \sigma^2 \text{ is UNKNOWN.}$$

If we replace σ^2 by S^2, what's the distribution of

$$T = \frac{\overline{X} - \mu}{\sqrt{S^2/n}}?$$

The random variable T defined above is called the *t-statistic*.

t-Statistic (2)

Dividing both the numerator and denominator of T by $\sqrt{\sigma^2/n}$, we can rewrite T as

$$T = \frac{(\overline{X} - \mu)\big/\sqrt{\sigma^2/n}}{\sqrt{S^2/\sigma^2}} = \frac{Z}{\sqrt{U/(n-1)}},$$

where

1. $Z = \frac{(\overline{X}-\mu)}{\sqrt{\sigma^2/n}} \sim N(0,1)$
2. $U = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$, and
3. Z and U are independent (from the indep of $\overline{X}$ and S^2).

From the definition of t-distribution, we know

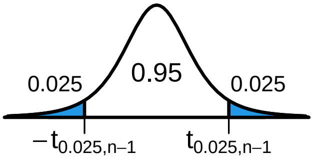
$$T = \frac{\overline{X} - \mu}{\sqrt{S^2/n}} \sim t_{n-1}$$

has a t-**distribution** with $n-1$ degrees of freedom.

95% One-Sample t-Confidence Interval

If $T \sim t_{n-1}$, let $t_{0.025,n-1}$ be the value so that

$$\mathrm{P}(-t_{0.025,n-1} \leq T \leq t_{0.025,n-1}) = 0.95$$



This means

$$\mathrm{P}\left(-t_{0.025,n-1} \leq T = \frac{\overline{X} - \mu}{\sqrt{S^2/n}} \leq t_{0.025,n-1} \right) = 0.95,$$

or equivalently

$$\mathrm{P}\left(\overline{X} - t_{0.025,n-1}\sqrt{\frac{S^2}{n}} \leq \mu \leq \overline{X} + t_{0.025,n-1}\sqrt{\frac{S^2}{n}} \right) = 0.95.$$

meaning, for 95% of the time, the sample mean $\overline{X}$ is within $t_{0.025,n-1}\sqrt{\frac{S^2}{n}}$ from the true value of μ.

95% One-Sample t-Confidence Interval

The interval

$$\left(\overline{X} - t_{0.025,n-1}\sqrt{\frac{S^2}{n}},\ \overline{X} + t_{0.025,n-1}\sqrt{\frac{S^2}{n}}\right).$$

is thus call the *95% one-sample t-confidence interval* for μ.

For example, with $n = 16$ observations, $t_{0.025,16-1} \approx 2.131$, the 95% confidence interval for μ is

$$\left(\overline{X} - 2.131\sqrt{\frac{S^2}{16}},\ \overline{X} + 2.131\sqrt{\frac{S^2}{16}}\right).$$