Sampling Distrbutions

Sampling Distributions

A sampling distribution is the distribution of a statistic.

Many known statistics have a known distribution.

$\bar{X}$

Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, show that $\bar{X} \sim N(\mu, \sigma^2/n)$.

Sum of χ_1^2

Let $Z_1^2, \dots, Z_n^2$ be a iid χ_1^2. Find $Y = \sum_{i=1}^n Z_i^2$

s^2

t-distribution

Let $Z \sim N(0,1)$, $W \sim \chi^2_\nu$, $Z \perp W$; therefore:

$$T = \frac{Z}{\sqrt{W/\nu}} \sim t_\nu$$

$X_1, \ldots, X_n \sim N(\mu, \sigma^2)$

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

HT: μ

is $\bar{X} \overset{?}{=} \mu$

given μ

$\bar{X}$ is known

μ is known

$\mu = 0$

if σ is unknown

$\mu = 1$

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \quad \sqrt{\frac{(n-1)S^2}{\sigma^2}/n-1}$$

$$\frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)s^2}{\sigma^2} / (n-1)}} \qquad \bar{X} \sim N(\mu, \sigma^2/n)$$

$$\frac{\frac{\bar{X}-\mu}{\sigma}}{\frac{s}{\sigma}} \rightarrow \frac{\bar{X}-\mu}{S/\sqrt{n}}$$

← from the data

$$t = \frac{\bar{X}-\mu}{S/\sqrt{n}}$$

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

$$Y \sim \chi^2(\nu)$$

t-distribution ↓

$$T = \frac{Z}{\sqrt{Y/\nu}} \sim t(\nu)$$

$$X \sim N(0,1) \quad Y \sim \chi^2(\nu)$$

$$T = \frac{X}{\sqrt{Y/\nu}} \sim t(\nu)$$

$X \perp Y$ independent

$\bar{X} \perp S^2$

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$f_Y(y) = \frac{1}{\Gamma(\nu/2)\, 2^{\nu/2}} y^{\nu/2 - 1} e^{-y/2}$$

$$T = \frac{X}{\sqrt{Y/\nu}} = X\sqrt{\nu/Y}$$

$W = Y \qquad T = X \cdot \sqrt{\nu/W}$

$$f_{WT}(w,t) = f_{XY}(x,y)\,|J|$$

Jacobian

determinant

$$f_{XY}(x,y) = f_X(x)\,f_Y(y)$$

$$= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\Gamma(\nu/2)\,2^{\nu/2}}\, y^{\nu/2-1}\, e^{-y/2}$$

$$T = X\sqrt{\nu/Y} = X\sqrt{\nu/W} \qquad Y = W$$

$$X = T\sqrt{W/\nu}$$

$$J = \begin{pmatrix} \frac{dx}{dT} & \frac{dx}{dw} \\ \frac{dy}{dT} & \frac{dy}{dw} \end{pmatrix} = \begin{pmatrix} \sqrt{w/v} & \frac{1}{2v\sqrt{w/v}} \\ 0 & 1 \end{pmatrix}$$

$$|J| = \left(\sqrt{w/v}\,(1) - \frac{1}{2v\sqrt{w/v}}(0) \right)$$

$$|J| = \sqrt{w/v}$$

$$f_{TW} = f_{XY}\,|J| = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \frac{1}{\Gamma(v/2)\,2^{v/2}} y^{v/2 - 1} e^{y/2} \sqrt{w/v}$$

$$f_{TW} = \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(T\sqrt{W/\nu}\right)^2}{2}} \frac{1}{\Gamma(\nu/2)2^{\nu/2}} \times$$

$$\times\ W^{\nu/2-1}\ e^{-W/2} \sqrt{W/\nu}$$

$$f_{TW} = \frac{1}{\sqrt{2\pi}\,\Gamma(\nu/2)2^{\nu/2}} W^{\nu/2-1}\ W^{1/2}\ \underline{\nu}^{-1/2}$$

$$\exp\left\{-\frac{T^2 W/\nu}{2} - W/2\right\}$$

$$= \frac{1}{\sqrt{2\pi\nu}\,\Gamma(\nu/2)2^{\nu/2}} W^{\nu/2-1/2}$$

$$\exp\left\{-\frac{\omega}{2}\left(\frac{T^2}{\nu}+1\right)\right\} = f_{T\omega}$$

$$f_T = \int_0^\infty f_{T\omega}\, d\omega$$

$$\int_0^\infty \frac{1}{\sqrt{2\pi\nu}\,\Gamma(\nu/2)2^{\nu/2}}\, \omega^{\nu/2-1/2} \exp\left\{-\omega/2\left(T^2/\nu+1\right)\right\} d\omega$$

$$k = \frac{1}{\sqrt{2\pi\nu}\,\Gamma(\nu/2)2^{\nu/2}}$$

$$K\int_0^\infty w^{\nu/2 - 1/2} \exp\{-w/2(T^2/\nu + 1)\} \, dw$$

$$w^x \exp\{-w y\}$$

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

shape $= \alpha$

rate $= \beta$

$$f_X = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-x\beta}$$

$$x^{\alpha-1} \leftarrow w^{\nu/2 - 1/2} \qquad \alpha = \frac{\nu+1}{2}$$

$x = w \qquad w^{\frac{\nu+1}{2} - 1}$

$$e^{-\frac{\omega}{2}\left(\frac{T^2}{\nu}+1\right)} \qquad e^{-x\beta}$$

$$\beta = \frac{1}{2}\left(\frac{T^2}{\nu}+1\right)$$

$$\int_0^\infty \frac{\beta^\alpha}{\Gamma(\nu/2)} x^{\alpha-1} e^{-x\beta} dx = 1$$

$$\int x^{\alpha-1} e^{-x\beta} dx = \frac{\Gamma(\alpha)}{\beta^\alpha}$$

$$K \int \omega^{\frac{\nu+1}{2}-1} e^{-\omega/2\left(\frac{T^2}{\nu}+1\right)} d\omega$$

$$k \frac{\Gamma\left(\frac{v+1}{2}\right)}{\left[\frac{1}{2}\left(\frac{T^2}{v}+1\right)\right]^{\frac{v+1}{2}}}$$

$$\frac{1}{\sqrt{2\pi v}\,\Gamma(v/2)\,2^{v/2}} \quad \frac{\Gamma\left(v+1/2\right)}{\left[\frac{1}{2}\left(\frac{T^2}{v}+1\right)\right]^{\frac{v+1}{2}}}$$

$$\frac{\Gamma\left(v+1/2\right)}{\Gamma(v/2)\,\sqrt{\pi v}\,2^{1/2}\,2^{v/2}\,2^{-(v+1/2)}} \left[\frac{T^2}{v}+1\right]^{-\frac{v+1}{2}}$$

$$\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{\pi v}}\left[\frac{T^2}{v}+1\right]^{-\frac{v+1}{2}} = f_T(t)$$

F-distribution

Let $W_1 \sim \chi^2_{\nu_1}$ $W_2 \sim \chi^2_{\nu_2}$, and $W_1 \perp W_2$; therefore:

$$F = \frac{W_1/\nu_1}{W_2/\nu_2} \sim F_{\nu_1,\nu_2}$$

Order Statistics

Order statistics are a fundamental concept in statistics and probability, dealing with the properties of sorted random variables. They provide insights into the distribution and behavior of sample data, such as minimum, maximum, and quantiles. Understanding order statistics is crucial in various fields such as risk management, quality control, and data analysis.

Order Statistics

Let $X_1, X_2, \ldots, X_n$ be a sample of n independent and identically distributed (i.i.d.) random variables with a common probability density function $f(x)$. The order statistics are the sorted values of this sample, denoted as:

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$$

Here, $X_{(1)}$ is the minimum, and $X_{(n)}$ is the maximum of the sample.

Order Statistics

- $X_{(k)}$: The k-th order statistic, representing the k-th smallest value in the sample.
- $X_{(1)}, X_{(n)}$: The minimum and maximum of the sample, respectively.

Distribution of Order Statistic

The distribution of the k-th order statistic $X_{(k)}$ can be derived using combinatorial arguments. Its PDF is given by:

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!}[F(x)]^{k-1}[1-F(x)]^{n-k}f(x$$

This formula shows how the distribution of $X_{(k)}$ depends on the underlying distribution of the sample and its position k.

Central Limit Theorem

Let $X_1, X_2, \dots, X_n$ be identical and independent distributed random variables with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$. We define

$$Y_n = \sqrt{n}\left(\frac{\bar{X} - \mu}{\sigma}\right) \text{ where } \bar{X} = x\frac{1}{n}\sum_{i=1}^{n} X_i.$$

Then, the distribution of the function Y_n converges to a standard normal distribution function as $n \to \infty$.

Central Limit Theorem

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

Central Limit Theorem Proof

Example

Let $X_1, \ldots, X_n \overset{iid}{\sim} \chi^2_p$, the MGF is $M(t) = (1 - 2t)^{-p/2}$. Find the distribution of $\bar{X}$ as $n \to \infty$.