MATH E-156 Mathematical Statistics

Harvard Extension School

Dmitry Kurochkin

Fall 2020
Lecture 8

Contents

Contents

Law of Large Numbers (LLN)

Thm. (LLN)
Let $X_1, X_2, X_3, \ldots$ be a sequence of independent random variables with

$$E(X_k) = \mu \quad \text{and} \quad \mathrm{Var}(X_k) = \sigma^2 \quad \text{for } k = 1, 2, 3, \ldots$$

Let

$$\bar{X}_n \doteq \frac{1}{n} \sum_{k=1}^{n} X_k.$$

Then, for any $\varepsilon > 0$,

$$P(|\bar{X}_n - \mu| > \varepsilon) \to 0 \quad \text{as} \quad n \to \infty.$$

Contents

Convergence in Distribution

Def.

A sequence of random variables $X_1, X_2, X_3, \ldots$ *converges in distribution* to a random variable X if

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x)$$

at every point at which $F_X(x)$ is continuous.
Here, $F_X(x)$ is the cdf of X and F_{X_n} is the cdf of X_n, $n = 1, 2, 3, \ldots$

Continuity Theorem

Thm. (Continuity Theorem)

Let $F_n(x)$, $n = 1, 2, 3, \ldots$ be a sequence of cumulative distribution functions with the corresponding moment generating functions $M_n(t)$.

Let $F(x)$ be a cumulative distribution function with the moment generating function $M(t)$.

If

$$M_n(t) \to M(t) \text{ as } n \to \infty$$

for all t in some neighborhood of 0, then

$$\lim_{n\to\infty} F_n(x) = F_X(x)$$

at every point at which $F_X(x)$ is continuous.

Central Limit Theorem (CLT)

Thm. (CLT)

Let $X_1, X_2, X_3, \dots$ be a sequence of independent identically distributed random variables with

$$E(X_k) = \mu \quad \text{and} \quad \mathrm{Var}(X_k) = \sigma^2 > 0 \quad \text{for } k = 1, 2, 3, \dots$$

Let

$$\bar{X}_n \doteq \frac{1}{n} \sum_{k=1}^{n} X_k.$$

Then, assuming mgf's of X_k exist in a neighborhood of 0,

$$\lim_{n \to \infty} P\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq x \right) = \Phi(x) \quad \text{for all} \quad x \in \mathbb{R},$$

where

$$\Phi(z) \doteq \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

is the cdf of a standard normal random variable.

Contents

Chi-Square Distribution

Def.

Let $Z \sim N(0,1)$. Then the distribution of $U = Z^2$ is called the *chi-square* distribution with 1 *degree of freedom*:

$$U \sim \chi_1^2.$$

Chi-Square Distribution

Def.

Let $Z \sim N(0,1)$. Then the distribution of $U = Z^2$ is called the *chi-square* distribution with 1 *degree of freedom*:

$$U \sim \chi_1^2.$$

Def.

Let $U_1, U_2, \ldots, U_n \overset{\text{iid}}{\sim} \chi_1^2$. Then the distribution of $V = U_1 + U_2 + \ldots + U_n$ is called the *chi-square* distribution with n *degrees of freedom*:

$$V \sim \chi_n^2.$$

Chi-Square Distribution

Claim

If $V \sim \chi^2_n$ then $V \sim \text{Gamma}\left(\underbrace{\frac{n}{2}}_{\alpha}, \underbrace{\frac{1}{2}}_{\lambda}\right)$, i.e. the pdf is

$$f_V(v) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} v^{(n/2)-1} e^{-v/2}, & \text{for all } v \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Chi-Square Distribution

Claim

If $V \sim \chi_n^2$ then $V \sim \text{Gamma}\Big(\underbrace{\frac{n}{2}}_{\alpha}, \underbrace{\frac{1}{2}}_{\lambda}\Big)$, i.e. the pdf is

$$f_V(v) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} v^{(n/2)-1} e^{-v/2}, & \text{for all } v \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Note:

If $V \sim \chi_n^2$ then the mgf is

$$M_V(t) = (1 - 2t)^{-n/2}.$$

Contents

t Distribution

Def.

Let $Z \sim N(0,1)$ and $U \sim \chi_n^2$ be independent. Then the distribution of $T = \frac{Z}{\sqrt{U/n}}$ is called the t distribution with n *degrees of freedom*:

$$T \sim t_n.$$

t Distribution

Def.
Let $Z \sim N(0,1)$ and $U \sim \chi_n^2$ be independent. Then the distribution of $T = \frac{Z}{\sqrt{U/n}}$ is called the t distribution with n *degrees of freedom*:

$$T \sim t_n.$$

Claim
If $T \sim t_n$ then the pdf of T is

$$f_T(t) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\,\Gamma(n/2)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \quad \text{for all } t \in \mathbb{R}.$$

Contents

F Distribution

Def.

Let $U \sim \chi^2_m$ and $V \sim \chi^2_n$ be independent. Then the distribution of $W = \frac{U/m}{V/n}$ is called the F distribution with m and n *degrees of freedom*:

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

F Distribution

Def.

Let $U \sim \chi^2_m$ and $V \sim \chi^2_n$ be independent. Then the distribution of $W = \frac{U/m}{V/n}$ is called the F distribution with m and n *degrees of freedom*:

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

Claim

If $W \sim F_{m,n}$ then the pdf of W is

$$f_W(w) = \begin{cases} \frac{\Gamma((m+n)/2)}{\Gamma(m/2)\Gamma(n/2)} \left(\frac{m}{n}\right)^{\frac{m}{2}} w^{\frac{m}{2}-1} \left(1 + \frac{m}{n}w\right)^{-\frac{m+n}{2}}, & \text{for all } w \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Contents

Definition of Sample Mean and Sample Variance

Def.

Let $X_1, X_2, \ldots, X_n$ be independent identically distributed (iid) random variables. Then we define:

- *sample mean* as

$$\bar{X} = \frac{1}{n} \sum_{k=1}^{n} X_k$$

- *sample variance* as

$$S^2 = \frac{1}{n-1} \sum_{k=1}^{n} \left(X_k - \bar{X} \right)^2$$

Contents

Distribution of Sample Mean

Claim

If $X_1, X_2, \ldots, X_n \stackrel{\text{iid}}{\sim} \text{Normal}(\mu, \sigma^2)$ then

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

Contents

Independence of Sample Mean and Sample Variance

Thm.

If $X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \text{Normal}(\mu, \sigma^2)$ then

$$\bar{X} \;\text{ and }\; (X_1 - \bar{X},\, X_2 - \bar{X}, \ldots, X_n - \bar{X}) \;\text{ are independent.}$$

Independence of Sample Mean and Sample Variance

Thm.

If $X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \text{Normal}(\mu, \sigma^2)$ then

$$\bar{X} \ \text{ and } \ (X_1 - \bar{X},\, X_2 - \bar{X}, \ldots,\, X_n - \bar{X}) \ \text{ are independent.}$$

Corollary

If $X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \text{Normal}(\mu, \sigma^2)$ then

$$\bar{X} \ \text{ and } \ S^2 \ \text{ are independent.}$$

Contents

Distribution of Sample Variance

Thm.

If $X_1, X_2, \ldots, X_n \stackrel{\text{iid}}{\sim} \text{Normal}(\mu, \sigma^2)$ then

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

Distribution of Sample Variance

Thm.

If $X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \text{Normal}(\mu, \sigma^2)$ then

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

Corollary

If $X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \text{Normal}(\mu, \sigma^2)$ then

$$\frac{\bar{X} - \mu}{S\sqrt{n}} \sim t_{n-1}.$$