Stat Inference | 1/23/25

Chapter 5

5.1) Basic Concepts of random samples

Def 5.1.1: The RVs $X_1,\dots,X_n$ are a random sample of size n from the population $f(x)$ if

1. $X_1, X_2,\dots,X_n$ are mutually independent RVs
2. The marginal pdf or pmf of each X_i the $\underbrace{\text{same function, } f(x)}_{\text{identically distributed}}$

$\Big\}$ iid

From Def 4.6.5, the joint pdf/pmf of $X_1,\dots,X_n$ is
$f(x_1,\dots,x_n) = f(x_1)\,f(x_2)\cdots f(x_n) = \prod f(x_i)$

If a population is a member of a random family, with pdf or pmf given by $f(x|\theta)$, then the joint pdf or pmf is:

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

with the same parameter value θ is used in each term in the product.

Ex: Normal $\quad \theta = \{\mu, \sigma^2\}$

$$f(x|\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2\sigma^2}\cdot(x-\mu)^2\right) \quad ; \begin{array}{l} x\in\mathbb{R} \\ \mu\in\mathbb{R} \\ \sigma^2>0 \end{array}$$

Joint pdf

$$f(x_1,\dots,x_n \mid \mu,\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2\sigma^2}\cdot(x_i-\mu)^2\right)$$

$$= \left(\frac{1}{\sigma 2\pi}\right)^{n/2} \cdot \exp\left(-\frac{1}{2\sigma^2}\cdot \Sigma(x_i-\mu)^2\right)$$

Ex: Poisson $\quad \theta = \{\lambda\}$

$$f(x|\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad ; \begin{array}{l} x = 0,1,\dots \\ \lambda \geq 0 \end{array}$$

joint pmf

$$f(x_1,\dots,x_n \mid \lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda}\,\lambda^{\Sigma x_i}}{\prod_{i=1}^{n} x_i!}$$

Subtle point: We are acting like we have an infinite population

In practice, we obtain $x_1, x_2, \dots, x_n$ sequentially

- First, the experiment is performed and $X_1 = x_1$ is observed.
- Second, the experiment is performed and $X_2 = x_2$ is observed.
- Independence in RVs implies that the dist'n of X_2 is unaffected by the fact that $X_1 = x_1$ was observed first. Or, formally, $f_{X_2}(x_2) = f_{X_2|X_1}(X_2 | X_1 = x_1)$

$$= \frac{f(x_1, x_2)}{f(x_1)} \; ; \; f(x_1) > 0$$

- Removing x_1 from the infinite population does not change the population, so $X_2 = x_2$ is still a random sample from the same population.

Def 5.1.1 may or may not apply to finite samples depending on how data collection is done.

Sampling with replacement from finite pop. satisfies Def 5.1.1.
Sampling w/o replacement does <u>not</u> satisfy Def 5.1.1.

In practice, not a big deal if we assume independence.

<u>Finite population example:</u> (w/o replacement)

$\mathcal{X} = \{x_1, \dots, x_n\}$. Let x and y be distinct elements of $\mathcal{X}$.

$P(X_2 = y \mid X_1 = y) = 0$ as y is gone.
$P(X_1 = x) = 1/n$
$P(X_2 = x) = 1/n = \sum P(X_2 = x \mid X_1 = x_i) \cdot P(X_1 = x_i)$

$$\hookrightarrow \begin{cases} = 0 & \text{for } x_i = x \\ = \frac{1}{n-1} & \text{for all others} \end{cases}$$

$$= 0\left(\frac{1}{n}\right) + (n-1)\left(\frac{1}{n-1} \cdot \frac{1}{n}\right)$$

$$= 1/n$$

↖ we get identically distributed but not independent.
∴ not a random sample.

For large population the difference is minimal.

<u>Example</u>: 5.1.3

$\mathcal{X} = \{1, 2, \dots, 1000\}$, $n = 10$ w/ replacement
$P(X_1 > 200, X_2 > 200, \dots, X_{10} > 200) = P(X_1 > 200) \cdot \ldots P(X_{10} > 200)$

$$= \left(\frac{800}{1000}\right)^{10} \approx 0.107$$

W/o replacement becomes a hypergeo. $= \dfrac{\binom{800}{10}\binom{200}{0}}{\binom{1000}{10}} \approx 0.1062$

approx the same

1/28/2025

5.2 Sums of RVs from random sample (RS)

Def 5.2.1: Let $X_1, \dots, X_n$ be a RS of size n from a population and let $T(X_1, \dots, X_n)$ be a real valued or vector valued function whose domain includes the sample space of $(X_1, \dots, X_n)$. Then the RV or RVector $T(X_1, \dots, X_n)$ is called a statistic. The prob dist'n of a statistic Y is called the sampling distribution of Y.

Statistic
a function that takes in data

Note: Statistics cannot be a function of unknown parameters.

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

Sample median, mode, max

Not a statistic:

$$Z_i = \frac{X_i - \mu}{\sigma}$$

we don't know these population values

Lemma 5.2.5) Let $X_1, \dots, X_n$ be a RS from a population and let $g(x)$ be a function s.t. $E[g(x)]$ and $Var[g(X)]$ exist $[-\infty < \text{value} < \infty]$. Then

$$E[\Sigma g(x_i)] = n E[g(X_1)]$$
$$Var[\Sigma g(x_i)] = n Var[g(X_1)]$$

Proof: $E(\Sigma g(x_i)) = \Sigma E[g(x_i)]$ (linearity of E)

$= \Sigma E[g(X_1)]$ (identically distributed)

$= n E[g(X_1)]$

$Var(\Sigma g(x_i)) = E\{[\Sigma g(x_i) - E[\Sigma g(x_i)]]^2\}$ (def of var)

$= E\{[\Sigma(g(x_i) - E[g(x_i)]]^2\}$ (linearity of E, simplifying)

$= E[\Sigma[g(x_i) - E(g(x_i)]]^2]$
$+ E[\sum_{1 \le i < j \le n} [g(x_i) - E(g(x_i))]\underbrace{[g(x_j) - E(g(x_j))]}_{Cov}]$

$= \Sigma E[g(X_i) - E[g(x_i)]]^2 + 2 \sum_{1 \le i < j \le n} Cov(g(x_i), g(X_j))$

$= \Sigma Var[g(x_i)] + 0$ (Cov is 0 due to ind)

$= \Sigma Var[g(X_1)]$ (identically distrib)

$= n Var[g(X_1)]$

Thm 5.2.6 : Let $X_1, \dots, X_n$ be a RS from a population with mean μ and variance $\sigma^2 < \infty$.

Then: a. $E[\bar{X}] = \mu$
b. $Var[\bar{X}] = \sigma^2/n$
c. $E[S^2] = \sigma^2$

Proof: a. $E[\bar{X}] = E[\frac{1}{n} \Sigma X_i]$

$= \frac{1}{n} E[\Sigma X_i]$

$= \frac{n}{n} \cdot \mu$ (lemma 5.2.5)

$= \mu$

b. $Var[\bar{X}] = Var(\frac{1}{n} \Sigma X_i)$

$= (\frac{1}{n})^2 Var(\Sigma X_i)$ (prop of var)

$= \frac{1}{n^2} \cdot n\sigma^2$ (lemma 5.2.5)

$= \frac{\sigma^2}{n}$

c. Note: $\Sigma (X_i - \bar{X})^2 = \Sigma (X_i^2 - 2\bar{X} X_i + \bar{X}^2)$
$= \Sigma X_i^2 - 2\bar{X} \Sigma X_i + n\bar{X}^2)$
$= \Sigma X_i^2 - 2n(\bar{X})^2 + n\bar{X}^2$
$= \Sigma X_i^2 - n(\bar{X})^2$

$E[S^2] = E\left[\frac{1}{n-1} \Sigma (X_i - \bar{X})^2\right]$ (def of samp var)

$= E\left[\frac{1}{n-1} (\Sigma X_i^2 - n(\bar{X})^2)\right]$ (w/ note)

$= \frac{1}{n-1} E[\Sigma X_i^2 - n(\bar{X})^2]$

$= \frac{1}{n-1} (\Sigma E[X_i^2] - n E[\bar{X}^2])$

$= \frac{1}{n-1} \left(\Sigma(\sigma^2 + \mu^2) - n\left(\frac{\sigma^2}{n} + \mu^2\right)\right)$

$= \frac{1}{n-1} \left(n\sigma^2 + n\mu^2 - \frac{n\sigma^2}{n} - n\mu^2\right)$

$= \frac{(n-1)\sigma^2}{(n-1)}$

$= \sigma^2$

Note:
$Var(X_i) = E[X_i^2] - (E(X_i)$
$E[X_i^2] = Var(X_i) + (E(X_i))^2$
$= \sigma^2 + \mu^2$
Similarly
$E[\bar{X}^2] = \frac{\sigma^2}{n} + \mu^2$

1/30/2025

Thm 5.2.11: Suppose $X_1, \dots, X_n$ is a RS from a pdf or pmf $f(x|\theta)$ where

$$f(x|\theta) = h(x)\, c(\theta) \exp\left(\sum w_i(\theta)\, t_i(x) \right)$$

Read the book my hands hurt

is a member of an exponential family, define statistics $T_1, \dots, T_K$ by

$$T_i(X_1, \dots, X_n) = \sum t_i(x_j), \quad i = 1, \dots, k$$

If the see... big picture: RS taken from exponential family are also from an exponential family

Specifically, statistics created from this RS

Example:

Let $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

$$f(\vec{x}\,|\theta) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2} (x_i - \mu)^2 \right)$$

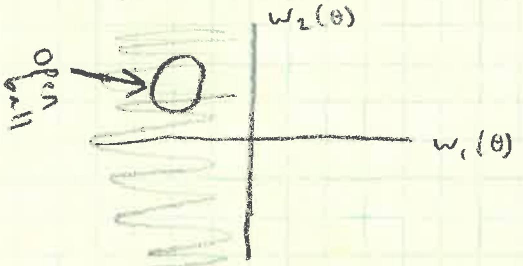
$$= (2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \sum (x_i - \mu)^2 \right)$$

$$= (2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2} \left(\sum x_i^2 - 2\mu \cdot \sum x_i + n\mu^2 \right) \right)$$

$$= \underbrace{(2\pi)^{-n/2}}_{h(\vec{x})} \underbrace{(\sigma^2)^{-n/2} \exp\left(\frac{-n\mu^2}{2\sigma^2} \right)}_{c(\theta)} \exp\left(\underbrace{-\frac{1}{2\sigma^2}}_{w_1(\theta)} \underbrace{\sum x_i}_{t_1(x)} + \underbrace{\frac{\mu}{\sigma^2}}_{w_2(\theta)} \underbrace{\sum x_i}_{t_2(x)} \right)$$

Note: $(w_1(\theta), w_2(\theta)) \in (-\infty, 0) \times (-\infty, \infty)$

Note: Something about can you draw an open circle in the space?



$\therefore$ Dist of $(\sum x_i^2, \sum x_i)$ is a member of an exponential family

5.3) Sampling from a Normal Dist leads to Statistics w/ many useful properties and well known sampling distributions.

5.3.1) Sampling dist of $\bar{X}$ and S^2

Theorem: If $X_1, \dots, X_n$ are ind of $Y_1, \dots, Y_n$, then $g_1(X_1, \dots, X_n)$ is ind of $g_2(Y_1, \dots, Y_2)$ where g_1 and g_2 are real valued functions.

Theorem: 5.3.1) Let $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$
Define $\bar{X} = \frac{1}{n} \Sigma X_i$, $S^2 = \frac{1}{n-1} \Sigma (X_i - \bar{X})^2$

Then
a) $\bar{X}$ and S^2 are ind RVs
b) $\bar{X} \sim N(\mu, \sigma^2/n)$
c) $\left(\frac{n-1}{\sigma^2}\right) \sim \chi^2_{n-1}$

$\nwarrow$ how? S^2 has $\bar{X}$ in its definition.

Proof: (of A)

$$\text{Let } \vec{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \sim N\Big(\overset{\vec{\mu}}{\begin{bmatrix} \mu \\ \vdots \\ \mu \end{bmatrix}}, \overset{\vec{\Sigma}}{\sigma^2 I_n}\Big)$$

$$\text{Let } A = \underbrace{\left.\begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \cdots & \frac{1}{n} \\ \frac{n-1}{n} & \frac{-1}{n} & \cdots & -\frac{1}{n} \\ \frac{-1}{n} & \frac{n-1}{n} & -\frac{1}{n} \cdots & -\frac{1}{n} \\ \vdots & & & \\ \frac{-1}{n} & \cdots & -\frac{1}{n} & \frac{n-1}{n} \end{bmatrix}\right\} n+1}_{n}$$

$$\text{Then } \vec{Y} = A\vec{X} = \underset{(n+1)+1}{\begin{bmatrix} \bar{X} \\ X_1 - \bar{X} \\ \vdots \\ X_n - \bar{X} \end{bmatrix}} \sim N(A\mu, A\Sigma A^t)$$

Linear transform of multi norm is also multi norm

$$A\mu = \left.\begin{bmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right\} n \text{ 0's}$$

$$\begin{aligned} &E(X_i - \bar{X}) \\ &= E[X_i] - E[\bar{X}] \\ &= \mu - \mu \\ &= 0 \end{aligned}$$

$$A\Sigma A^T = [\quad] \begin{bmatrix} \sigma^2 & & & 0 \\ & \sigma^2 & & \\ & & \sigma^2 & \\ 0 & & & \sigma^2 \end{bmatrix} \begin{bmatrix} \frac{1}{n} & \frac{n-1}{n} & \\ \frac{1}{n} & -\frac{1}{n} & \\ \vdots & \vdots & \cdots \\ \frac{1}{n} & -\frac{1}{n} & \end{bmatrix} = \begin{bmatrix} \sigma^2/n & \sigma^2/n & \sigma^2/n \\ \frac{n-1}{n}\sigma^2 & -\frac{1}{n}\sigma^2 & -\frac{1}{n}\sigma^2 \\ & & \end{bmatrix} \begin{bmatrix} \quad \end{bmatrix}$$

$$= \begin{bmatrix} \frac{n\sigma^2}{n^2} & \frac{\sigma^2}{n}\left(\frac{n}{n} - \frac{1}{n} - \frac{1}{n}\right. \\ & \end{bmatrix}$$

Oh god
Out of space

⑥

Proof Ctn

$$A\Sigma A^T = \begin{bmatrix} \overset{[1,1]}{\downarrow} & \overset{\text{element } [1,2]}{\downarrow} & \\ \frac{n\sigma^2}{n^2} & \frac{\sigma^2}{n}\left(\overbrace{\frac{n}{n} - \frac{1}{n} - \frac{1}{n} - \cdots - \frac{1}{n}}^{\text{so } \frac{n}{n} \text{ and we have } \frac{n}{n} - \frac{n}{n} = 0}\right) & \cdots\cdots \\ \frac{1}{n}\left(\sigma^2\left(\frac{n}{n} \underbrace{- \frac{1}{n} - \cdots - \frac{1}{n}}_{n/n \ \uparrow \text{ same thing here}}\right)\right) & \cdots & \\ \uparrow & & \\ \text{element } [2,1] & \vdots & \end{bmatrix} = \begin{bmatrix} \sigma^2/n & 0 \cdots 0 \\ 0 & \\ \vdots & \boxed{? \ \text{unimportant}} \\ 0 & \end{bmatrix}$$

recall that $\boxed{\text{normal}}$ rvs are $\boxed{\text{ind}}$ <u>if</u> their $\boxed{\text{covariance is zero}}$

From our work, $\text{Cov}(\bar{X}, X_1 - \bar{X}) = 0$

$\text{Cov}(\bar{X}, X_2 - \bar{X}) = 0$

$\vdots$

$\text{Cov}(\bar{X}, X_n - \bar{X}) = 0$

and

$\bar{X}, X_1 - \bar{X}, \dots, X_n - \bar{X}$ are all normal

So $\bar{X}$ is ind of $(X_1 - \bar{X}, \dots, X_n - \bar{X})$

$$S^2 = \sum \frac{(X_i - \bar{X})^2}{n-1} = g(X_1 - \bar{X}, \dots, X_n - \bar{X})$$

$\therefore$ $\bar{X}$ and S^2 are ind using thm before 5.3.1

ⓑ Dist of $\bar{X}$ should $\sim N(\mu, \sigma^2/n)$

Using mgfs $E[e^{t\bar{X}}]$ Recall $M_Y(t) = E[e^{tY}]$

If $Y \sim N(\mu, \sigma^2)$ then its mgf $M_Y(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$ ✪

↖ Known from back of book

$$M_{\bar{X}}(t) = E[e^{t\bar{X}}] = E[e^{t\frac{1}{n}\sum X_i}]$$

$$= E[e^{\frac{t}{n}X_1} e^{\frac{t}{n}X_2} \cdots e^{\frac{t}{n}X_n}]$$

$$= \prod_{i=1}^{n} \underbrace{E[e^{\frac{t}{n}X_i}]}_{M_{X_i}(\frac{t}{n})}$$

(since X_i's are ind, $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$)

$$= \prod_{i=1}^{n} \exp\left(\mu \frac{t}{n} + \sigma^2 \frac{t^2}{2n^2}\right)$$

(using ✪)

(part c is homework vs mgf method)

$$= \exp\left(n\mu \frac{t}{n} + \frac{n\sigma^2}{n^2}\frac{t^2}{2}\right)$$

$$= \underbrace{\exp\left(\mu t + \frac{\sigma^2}{n}\frac{t^2}{2}\right)}_{\text{mgf of a } N(\mu, \frac{\sigma^2}{n})} \quad \therefore \bar{X} \sim N(\mu, \sigma^2/n)$$

5.3.1 Proof Ctn

ⓒ $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

$\bar{X} = \frac{1}{n} \Sigma X_i$, $S^2 = \frac{1}{n-1} \Sigma (X_i - \bar{X})^2$

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

Note: $\frac{(n-1)S^2}{\sigma^2} = \frac{\Sigma (X_i - \bar{X})^2}{\sigma^2}$

$= \left(\frac{\Sigma (X_i - \bar{X})}{\sigma}\right)^2$

Also: $\underbrace{\sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma}\right)^2}_{\chi^2_{df=n}} = \Sigma \left[\frac{(X_i - \bar{X})}{\sigma} + \frac{(\bar{X} - \mu)}{\sigma}\right]^2$

(adding $0 = -\frac{\bar{X}}{\sigma} + \frac{\bar{X}}{\sigma}$)

$\Sigma \left[\frac{(X_i - \bar{X})}{\sigma}\right]^2 + \Sigma \left[\frac{\bar{X} - \mu}{\sigma}\right]^2 + 2\frac{(\bar{X} - \mu)}{\sigma^2} \underbrace{\Sigma (X_i - \bar{X})}$

(square and dist sum)

Note $\Sigma (X_i - \bar{X}) = 0$

b/c

$\Sigma X_i - n\bar{X}$

$= \Sigma X_i - n\frac{\Sigma X_i}{n}$

$= \Sigma X_i - \Sigma X_i = 0$

$= \Sigma \left[\frac{(X_i - \bar{X})}{\sigma}\right]^2 + n\left(\frac{\bar{X} - \mu}{\sigma}\right)^2$

$= \underbrace{\Sigma \left[\frac{(X_i - \bar{X})}{\sigma}\right]^2}_{\substack{\uparrow \\ V}} + \underbrace{\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}\right)^2}_{\substack{\uparrow \\ W}}$

Let

and $\Sigma \left(\frac{X_i - \mu}{\sigma}\right)^2$ be U

We know $U \sim \chi^2_n$

and $W \sim \chi^2_1 = Z^2$

recall $\bar{X} \sim N(\mu, \sigma^2/n)$

so $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

From our previous derivation

$M_U(t) = M_{V+W}(t)$ (def of mgf)

$= E[e^{t(V+W)}]$

$= E[e^{tV}]E[e^{tW}]$ (since W, V are ind by 5.3.1 a)

$(1-2t)^{-\frac{n}{2}} = E[e^{tV}] \cdot (1-2t)^{-\frac{1}{2}}$

$E[e^{tV}] = (1-2t)^{-\frac{n-1}{2}}$

$M_V(t) = (1-2t)^{-\frac{n-1}{2}}$ = mgf of χ^2_{n-1}

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

7.5

Lemma 5.3.2 χ^2 RVs

a. If $Z \sim N(0,1)$ then $Z^2 \sim \chi^2_1$

b. If $X_1, \ldots, X_n$ are independent w/ p_i df, respectively ($X_i \sim \chi^2_{p_i}$)
then $\Sigma X_i \sim \chi^2_{p_1 + \ldots + p_n}$

Proof: a) $M_{Z^2}(t) = E[e^{tZ^2}]$

$$= \int e^{tZ^2} \cdot \frac{e^{-Z^2/2}}{\sqrt{2\pi}} \Big\} \text{pdf of } N(0,1)$$

seems to have $\mu = 0$, $\sigma^2 = (1-2t)$

$$= \int \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{Z^2}{2}(1-2t)\right) dZ \qquad \Bigg| \qquad f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} y^2\right)$$

$$= \underbrace{\frac{1}{\sqrt{1-2t}} \int \frac{\sqrt{1-2t}}{\sqrt{2\pi}}}_{\text{multiply by "1"}} \exp\left(-\frac{(1-2t)}{2} z^2\right) dz$$

rewrite exp

$$\underbrace{\frac{\sqrt{1-2t}}{\sqrt{2\pi}} \exp\left(-\frac{(1-2t)}{2} z^2\right)}_{\text{pdf of } N(0, (1-2t)^{-1})}$$

$$= (1-2t)^{-1/2}$$
mgf of a χ^2

Proof B) $M_{\Sigma X_i}(t) = E[e^{t \Sigma X_i}]$

$$= E\left[\prod_{i=1}^{n} e^{tX_i}\right] \quad \text{(properties of exponent)}$$

$$= \prod E[e^{tX_i}] \quad (X_i\text{'s are ind})$$

$$= \prod M_{X_i}(t)$$

$$= \prod_{i=1}^{n} (1-2t)^{-p_i/2}$$

$$= (1-2t)^{-\frac{1}{2}(p_1 + \ldots + p_n)}$$

(mgf of a χ^2 w/ $p = p_1 + \ldots + p_n$)

Def 5.3.4) Recall: $\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right) \sim N(0,1)$

Let $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

Then $\frac{(\bar{X}-\mu)}{S/\sqrt{n}} \sim t_{n-1}$

A T random variable w/ p degrees of freedom has pdf

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

Proof: $\frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \cdot \frac{\frac{1}{\sigma/\sqrt{n}}}{\frac{1}{\sigma/\sqrt{n}}}$

$$= \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\frac{S}{\sigma} \cdot \frac{1/\sqrt{n}}{1/\sqrt{n}}}$$

$$= \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{S^2}{\sigma^2}}} = \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2}{\sigma^2} \cdot \frac{1}{(n-1)}}} = \frac{u}{\sqrt{V/p}}$$

w/ $u \sim N(0,1)$ and $V \sim \chi^2_{n-1}$

u, V independent

Now we look at $f_{u,v}(u,v)$

$$f_{u,v}(u,v) = \underbrace{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right)}_{\text{pdf of } N(0,1)} \cdot \underbrace{\frac{1}{\Gamma\left(\frac{p}{2}\right) 2^{p/2}} \cdot V^{(p/2)-1} e^{-V/2}}_{\text{pdf of } \chi^2_p \text{ where } p=n-1}$$

Let $t = \frac{u}{\sqrt{V/p}}$, $u = t\sqrt{V/p} = t\sqrt{w/p}$, $w = V$

$$J = \begin{vmatrix} \frac{du}{dt} & \frac{du}{dw} \\ \frac{dv}{dt} & \frac{dv}{dw} \end{vmatrix} = \begin{vmatrix} \sqrt{\frac{w}{p}} & \frac{t}{\sqrt{p}}\left(\frac{1}{2} w^{-1/2}\right) \\ 0 & 1 \end{vmatrix} = \left(\frac{w}{p}\right)^{1/2}$$

$$f_T(t) = \int_0^\infty f_{u,v}\left(t\left(\frac{w}{p}\right)^{\frac{1}{2}}, w\right) \cdot \left(\frac{w}{p}\right)^{\frac{1}{2}} dw$$

$$= \frac{1}{(2\pi)^{1/2} \Gamma\left(\frac{p}{2}\right) 2^{p/2}} \int_0^\infty e^{-(1/2)t^2 w/p} \cdot w^{(p/2)-1} \cdot e^{-w/2} \cdot \left(\frac{w}{p}\right)^{1/2} \cdot dw$$

$= \text{"}$ $\int_0^d$ Stuff that fits the kernel of a gamma$\left(\frac{p+1}{2}, \frac{2}{(1+t^2/p)}\right)$ pdf

2/6/2025

F distribution: 5.3.6 (a type of variance ratio distribution)

The ratios $\frac{S_X^2}{\sigma_X^2}$ and $\frac{S_Y^2}{\sigma_Y^2}$ are each χ^2 variates and $X \perp\!\!\!\perp Y$

so $F =$

key notes in the book

Note: the F is the ratio of two χ^2 RVs divided by their degrees of freedom.

$$F = \frac{\frac{(n-1) S_X^2}{(n-1)\sigma_X^2}}{\frac{(m-1) S_Y^2}{(m-1)\sigma_Y^2}} = \frac{\frac{u_1^2}{n-1}}{\frac{u_2^2}{m-1}}$$

where $u_1 \sim \chi^2_{n-1}$ and

$u_2 \sim \chi^2_{m-1}$ $\quad u_1 \perp\!\!\!\perp u_2$

Thm. 5.3.8

a) If $X \sim F_{p,q}$ then $\frac{1}{X} \sim F_{q,p}$ (flip the ratio, flip the parameters)

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

explan: $t_q = \frac{Z}{\sqrt{u_q/q}}$ where $Z \sim N(0,1)$, $u \sim \chi^2_q$ so /

$$\text{so } t_q^2 = \frac{Z^2 \leftarrow \chi^2_1}{u_q/q \leftarrow \chi^2_q}$$

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

5.4 Order Statistics

Def 5.4.1 order stats of a sample are the values placed in asc order.

Notation: $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ and

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

$$X_{(1)} = \min\{X_1, \ldots, X_n\}$$
$$X_{(n)} = \max\{X_1, \ldots, X_n\}$$
$$X_{(2)} = \text{2nd smallest } X_i$$

Sample range is $R = X_{(n)} - X_{(1)}$ Note! different from other common range $[X_{(1)}, X_{(n)}]$

Sample median is:

$$M = \begin{cases} X_{((n+1)/2)} & \text{if } n \text{ is odd} \\ \left(X_{(n/2)} + X_{(n/2)+1}\right)/2 & \text{if } n \text{ is even} \end{cases}$$

Thm 5.4.4 Let $X_{(1)}, \ldots, X_{(n)}$ denote the order statistics of a RS $X_1, \ldots, X_n$ from a continuous population w/ cdf $F_X(x)$ and pdf $f_X(x)$. Then the pdf of $X_{(j)}$ is

$$f_{X_{(j)}}(x) = \underbrace{\frac{n!}{(j-1)!\,(n-j)!}}_{\binom{n}{j-1}} \underbrace{f_X(x)}_{\text{density for } f_X(x)} [F_X(x)]^{j-1} [1 - F_X(x)]^{n-j}$$

Note! This is only for continuous as ties break this. For cont. $P(X_i == X_j) = 0$

Example! Exponential (β)

$$f(x|\beta) = 1/\beta \, \exp\left(-\frac{x}{\beta}\right), \quad \beta > 0, \quad 0 \leq x < \infty$$
$$F_X(x|\beta) = 1 - e^{-x/\beta}$$

What is the dist'n of the median if n is odd

$$f_{X\left(\frac{n+1}{2}\right)}(x) = \frac{n!}{\left(\frac{n+1}{2} - 1\right)!\left(n - \frac{n+1}{2}\right)!} \frac{1}{\beta} \exp\left(-\frac{x}{\beta}\right) \left[1 - e^{-x/\beta}\right]^{\frac{n+1}{2} - 1} \left[e^{-x/\beta}\right]^{n - \frac{n+1}{2}}$$

$0 \leq x < \infty$

$\beta > 0$

Thm 5.4.6 Same setup as 5.4.4

The joint pdf of $X_{(i)}$ and $X_{(j)}$, $1 \leq i < j \leq n$ is

$$f_{X_{(i)}X_{(j)}}(u,v) = \frac{n!}{(i-1)!(j-1-i)!(n-j)!} f_X(u) f_X(v) [F_X(u)]^{i-1} \nwarrow \text{cdf below ;}$$

$$\cdot \underset{\nwarrow \text{ cdf between}}{[F_X(v) - F_X(u)]^{j-1-i}} \; \underset{\nwarrow \text{ cdf above}}{[1 - F_X(v)]^{n-j}}$$

for $-\infty < u < v < \infty$

$(i-1)!$: values less than i
$(n-j)!$: values greater than j
$(j-1-i)!$: values between

$$\underbrace{X_{(1)} \quad X_{(2)}}_{2 \text{ less}} \quad \boxed{X_{(3)}} \quad \underbrace{X_{(4)} \quad X_{(5)}}_{2 \text{ between}} \quad \boxed{X_{(6)}} \quad \underbrace{X_{(7)} \quad X_{(8)}}_{2 \text{ above}}$$

joint density of all n order statistics:

$$f_{X_{(1)},\dots,X_{(n)}}(x_1,\dots,x_n) = \begin{cases} n!\, f_X(x_1) f_X(x_2) \cdots f_X(x_n) & -\infty < x_1 < x_2 < \cdots < x_n < \infty \\ 0 & \text{o/w} \end{cases}$$

Cause all the pdfs and cdfs result in 1 everywhere so it simplifies down

Example: $X_1,\dots,X_n \overset{iid}{\sim}$ w/ density $f(s)$

Define $R = X_{(n)} - X_{(1)}$ and $M = \dfrac{X_{(1)} + X_{(n)}}{2}$ | Find joint dist of (R, M)

This is a function of $X_{(1)}$ and $X_{(n)}$, so lets find $f_{X_{(1)}X_{(n)}}(u,v)$

$$f_{X_{(1)}X_{(n)}}(u,v) = \begin{cases} \dfrac{n!}{\underset{\text{below}}{0!}\,\underset{\text{everything is between}}{(n-2)!}\,\underset{\text{above}}{0!}} (F(v) - F(u))^{n-2} \cdot f(v) f(u) & u < v \\ 0 & \text{o/w} \end{cases}$$

$$= n(n-1)\,[F(v) - F(u)]^{n-2} f(u) f(v)$$

Note: $u = X_{(1)} = M - R/2$
$v = X_{(n)} = M + R/2$
(write $X_{(1)}, X_{(n)}$ as functions of M and R)

ex continued:

$$|J| = \det \begin{bmatrix} \frac{du}{dm} & \frac{du}{dR} \\ \frac{dv}{dm} & \frac{dv}{dR} \end{bmatrix} = \det \begin{bmatrix} 1 & -1/2 \\ 1 & 1/2 \end{bmatrix} = \frac{1}{2} - \left(-\frac{1}{2}\right) = 1$$

$$f_{M,R}(m,r) = f_{X_{(1)}X_{(2)}}\left(m - r/2 \; , \; m + r/2\right) \cdot 1$$

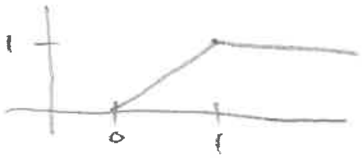
$$= n(n-1)\left[F\left(m + \frac{r}{2}\right) - F\left(m - \frac{r}{2}\right)\right]^{n-2} \cdot f\left(m - \frac{r}{2}\right) f\left(m + \frac{r}{2}\right) I_{\underset{\uparrow}{(0,\infty)}}($$

r must be pos

Feb 11 / 2025
Stat inf

Ex 5.4.7) Dist'n of the midrange and range for iid $U(0,1)$

If $X_1, \dots, X_n$ iid $U(0,1)$, then
$f(u) = I_{(0,1)}(u)$
$F(u) = u\, I_{(0,1)}(u) + I_{[1,\infty)}(u)$



$\int_0^u 1\, dv$
$u\, I_{(0,1)}(u)$

Range: $R = X_{(n)} - X_{(1)}$
Midrange: $M = (X_{(1)} + X_{(n)})/2$

$f_{M,R}(m,r) = n(n-1)\left[F(n+r/2) - F(m-r/2)\right]^{n-1} f(m-r/2) f(m+r/2)\, I_{(0,\infty)}(r)$ (from previous notes)

applying this to our example

$f_{M,R}(M,r) = n(n-1)\, I_{(0,1)}(m+r/2)\, I_{(0,1)}(m-r/2)\, I_{(0,\infty)}(r) \left[m+\frac{r}{2} - (m-\frac{r}{2})\right]^{n-2}$

$= n(n-1)\, I_{(0,1)}(m+\frac{r}{2})\, I_{(0,1)}(m-\frac{r}{2})\, I_{(0,\infty)}(r)\, [r]^{n-2}$ ★ $\nwarrow$ $F(m+\frac{r}{2}) = m+\frac{r}{2}$
$F(m-\frac{r}{2}) = m-\frac{r}{2}$

Note: $0 < m+\frac{r}{2} < 1 \Rightarrow -\frac{r}{2} < m < 1-\frac{r}{2}$
$0 < m-\frac{r}{2} < 1 \Rightarrow \frac{r}{2} < m < 1+\frac{r}{2}$
and $0 < r$
$\Rightarrow \frac{r}{2} < m < 1-\frac{r}{2}$

and, since $X_{(n)} \leq 1$ and $X_{(1)} \geq 0$, $R = X_{(n)} - X_{(1)} \leq 1$

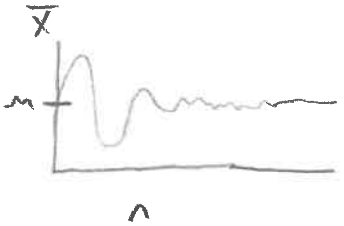
from ★ we have
$f_{M,R}(m,r) = n(n-1)\, I_{(\frac{r}{2}, 1-\frac{r}{2})}(m)\, I_{(0,1)}(r)\, [r]^{n-2}$

5.5 Convergence Concepts

Def 5.5.1) A seq. of RVs $X_1, \dots, X_n$ converges in prob to a RV, X, if, for every $\epsilon > 0$,
$\lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0$ or, equivalently, $\lim_{n\to\infty} P(|X_n - X| < \epsilon) = 1$

Markov's Inequality: If X is a non-negative RV, and $a > 0$,
then $P(X \geq a) \leq \frac{E[X]}{a}$

Thm 5.5.2 (Weak law of large numbers):
Let $X_1, \dots, X_n$ w/ $E[X_i] = \mu$, and $V(X_i) = \sigma^2$
Define $\bar{X}_n = \frac{1}{n} \Sigma X_i$, Then, for every $\epsilon > 0$
$\lim_{n\to\infty} P(|\bar{X}_n - \mu| < \epsilon) = 1$



Proof: $P(|\bar{X}_n - \mu| \geq \epsilon) = P((X_n - \bar{X})^2 \geq \epsilon^2) \leq \frac{E[(\bar{X}_n - \mu)^2]}{\epsilon^2}$ ↙ Thm 5.2.6b

$= \frac{Var(\bar{X}_n)}{\epsilon^2} = \frac{\sigma^2/n}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}$

$\to 0$ as $n \to \infty$

$\therefore 0 \leq \lim_{n\to\infty} P(|\bar{X}_n - \mu| > \epsilon) \leq 0$
$0 \leq x \leq 0$ means $x = 0$ so $P(" ") = 0$
$\therefore \bar{X}_n \xrightarrow{P} \mu$

Thm 5.5.4) Suppose $X_1, \dots, Y_n$ converges in probability to a RV X, and that h is a continuous function. Then $h(X_1), h(X_2), \dots$ converges in prob to $h(X)$.
More simply
if $X_n \xrightarrow{P} X$ then $h(X_n) \xrightarrow{P} h(X)$ if h is cont.

Def of joint convergence in prob:

$\underbrace{(X_n, Y_n)}_{\text{vector}} \xrightarrow{P} (X, Y)$ if

$$\lim_{n \to \infty} P\left(\sqrt{(X_n - X)^2 + (Y_n - Y)^2 \geq \epsilon}\right) = 0$$

The euclidean norm of a vector $X = (X_1, X_2)$ is $||X|| = \sqrt{X_1^2 + X_2^2}$

$$\therefore \sqrt{(X_n - X)^2 + (Y_n - Y)^2} = \underbrace{||(X_n, Y_n) - (X, Y)||}_{\text{diff in vectors}} = ||(X_n - X) + (Y_n - Y)||$$

By the triangle inequality

$$||(X_n - X) + (Y_n - Y)|| \leq ||X_n - X|| + ||Y_n - Y||$$

Thm: If $X_n \xrightarrow{P} X$ and $Y_n \xrightarrow{P} Y$, $(X_n, Y_n) \xrightarrow{P} (X, Y)$

$$\lim_{n \to \infty} P(||(X_n, Y_n) - (X, Y)|| \geq \epsilon) \leq \lim_{n \to \infty} \underbrace{P(||X_n - X|| + ||Y_n - Y|| \geq \epsilon)}_{\text{This implies } ||X_n - X|| \text{ or } ||Y_n - Y|| > \epsilon/2} \quad (\text{from } \Delta \text{ ineq})$$

$$\leq \lim_{n \to \infty} P\left(||X_n - X|| \geq \tfrac{\epsilon}{2}\right) + P\left(||Y_n - Y|| \geq \tfrac{\epsilon}{2}\right)$$

$= 0 + 0$ since $X_n \xrightarrow{P} X$ and $Y_n \xrightarrow{P} Y$

$\therefore (X_n, Y_n) \xrightarrow{P} (X, Y)$

Multivariate Continuous Mapping Theorem:

Let $\vec{X}_n$ be a sequence of K-dimensional random vectors

Let $g: \mathbb{R}^K \to \mathbb{R}^j$ be a continuous function. Then

$$\vec{X}_n \xrightarrow{P} X \implies g(\vec{X}_n) \xrightarrow{P} g(X)$$

and

$$\vec{X}_n \xrightarrow{D} X \implies g(\vec{X}_n) \xrightarrow{D} g(X)$$

Note:

$$\vec{X}_n = \begin{bmatrix} X_{1,n} \\ X_{2,n} \\ X_{3,n} \\ \vdots \\ X_{K,n} \end{bmatrix}$$

explanation:

This is a vector of k rvs that depend on n.

So, Say we have $X_n \sim N(0, \sigma^2/n)$

$$\vec{X}_1 = \begin{bmatrix} X_{1,n=1} \\ X_{2,n=1} \\ X_{3,n=1} \\ \vdots \\ X_{K,n=1} \end{bmatrix}$$

This stuff isnt in the book but is important to some examples

Application:

If $\underbrace{X_n \xrightarrow{P} X \text{ and } Y_n \xrightarrow{P} Y}_{\text{Univariate}}$

then $X_n + Y_n \xrightarrow{P} X + Y$
$X_n Y_n \xrightarrow{P} XY$

Example 5.5.3 Consistency of S^2

$X_1, \ldots, X_n$ iid F w/ mean μ and variance $\sigma^2 < \infty$

$$S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 = \frac{1}{n-1}\left[\sum X_i^2 - n(\bar{X}_n)^2\right] \Leftarrow (\text{Previous results})^{\text{Thm } 5.2.6}$$

Intermission: WLLN recap:

Let $X_1, \ldots, X_n \overset{iid}{\sim} F$ w/ $E[X_i] = \mu$ and $Var[X_i] = \sigma^2 < \infty$
Define $\bar{X}_n = \frac{1}{n}\sum X_i$ Then $\bar{X}_n \xrightarrow{P} \mu$

$$= \frac{n}{n-1} \cdot \left[\frac{\sum X_i^2}{n} - \frac{n}{n}(\bar{X}_n)^2\right] \quad (\text{mult by } \tfrac{n}{n} = 1)$$

$$\begin{array}{lcccc} & \frac{n}{n-1} & \frac{\sum X_i^2}{n} & \frac{n}{n} & (\bar{X}_n)^2 \\ (\text{converges to}) & \downarrow & \downarrow P ☆ & \| & \downarrow P \\ & [1] & [\sigma^2 + \mu^2] & 1 & \mu^2 \end{array}$$

$$☆\left(\begin{array}{l} \sigma^2 = E[X_i^2] - \mu^2 \\ E[X_i^2] = \sigma^2 + \mu^2 \end{array}\right)$$

$$= 1 \cdot [\sigma^2 + \mu^2 - \mu^2]$$

$$= \sigma^2$$

∴ The sample variance converges in probability to the population variance

5.5.5

Example: Consistency of S

Same setup as example 5.5.3

$S_n \xrightarrow{P} \sigma^2$

$h(x) \to \sqrt{x}$ is a continuous function

$\therefore \sqrt{S_n^2} = S_n \xrightarrow{P} \sqrt{\sigma^2} = \sigma$ by Thm 5.5.4

Example:

Let $X_1, \dots, X_n \overset{iid}{\sim} (0,1)$

Show $X_{(n)} \xrightarrow{P} 1$ (Max of the sample converges to 1)

Let $0 < \epsilon \leq 1$

Then $P(|X_{(n)} - 1| \geq \epsilon)$ (def. of convergence in prob)

$= P(|1 - X_{(n)}| \geq \epsilon)$ (property of abs)

$= P(1 - X_{(n)} \geq \epsilon)$ (remove abs as $X_{(n)} \leq 1$)

$= P(X_{(n)} \leq 1 - \epsilon)$ (good ol' algebra)

$= F_{X_{(n)}}(1 - \epsilon)$ (def of cdf)

$= \left(\prod_{i=1}^{n} P(X_i \leq 1 - \epsilon)\right)$ (since all n rvs $\leq 1 - \epsilon$ if $X_{(n)} \leq 1 - \epsilon$)

$= (1 - \epsilon)^n$ (X_i is $U(0,1) \to P(X_i \leq 1 - \epsilon) = \int_0^{1-t} 1 \, du$

$= 0$ as $n \to$ since $0 \leq 1 - \epsilon < 1$ (like multiplying 0.25 over and over again)

$\therefore X_{(n)} \xrightarrow{P} 1$

Convergence in distribution

Def 5.5.10 A sequence of r.vs. $X_1, \dots, X_n$ converges in distribution to a random variable X if

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

at all points x for which $F_X(x)$ is continuous

Example: 5.5.11

$X_1, \dots, X_n \overset{iid}{\sim} U(0,1)$

basically a lemma to use later

$$F_{X_{(n)}}\left(1-\frac{t}{n}\right) = P\left(X_{(n)} \le 1-\frac{t}{n}\right)$$

$$= P\left(X_1 \le 1-\frac{t}{n}, X_2 \le 1-\frac{t}{n}, \dots, X_n \le 1-\frac{t}{n}\right)$$

$$= \prod_{i=1}^{n} P\left(X_1 \le 1-\frac{t}{n}\right) \quad (X_i\text{'s are iid})$$

$$= \left(1-\frac{t}{n}\right)^n \quad (\text{eval cdf of } X_1 \text{ at } 1-\tfrac{t}{n})$$

$$\to e^{-t} \quad (\text{by def of exponential function})$$

$$P(\underbrace{n(1-X_{(n)})}_{\text{defines a sequence of rv's.}} \le t)$$

$$= P\left(-X_{(n)} \le -1+\frac{t}{n}\right) \quad (\text{algebra})$$

$$= P\left(1-\frac{t}{n} \le X_{(n)}\right) \quad (\text{algebra})$$

$$= 1-P\left(X_{(n)} \le 1-\frac{t}{n}\right) \quad (\text{by compliment rule and } X_{(n)} \text{ is cont.})$$

$$\to 1-e^{-t} \quad (\text{by previous work})$$

Note that $1-e^{-t}$ is cdf of $\text{Exp}(1)$ rv.

$\therefore n(1-X_{(n)}) \xrightarrow{P} \text{Exp}(1)$

Thm 5.5.12

If a sequence of random variables converges in probability to a random variable X, then the sequence also converges in distribution to X.

$$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{D} X$$

ie. convergence in probability is a stronger form of convergence than distribution

2/18/2025

Theorem: If $X_n \xrightarrow{D} c \in \mathbb{R}$, then $X_n \xrightarrow{P} c$

Since $X_n \xrightarrow{D} c$, $\forall \epsilon > 0$,

$$\lim_{n \to \infty} F_{X_n}(c-\epsilon) \qquad \left(F_c(x) = I_{[c,\infty)}(x) \right)$$

$$= F_c(c-\epsilon) = I_{[c,\infty)}(c-\epsilon) = 0$$

$$\lim_{n \to \infty} F_{X_n}\left(c + \frac{\epsilon}{2}\right) = F_c\left(c + \frac{\epsilon}{2}\right) = I_{[c,\infty)}\left(c + \frac{\epsilon}{2}\right) = 1$$

Note

$X_n \xrightarrow{D} X$ if

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

Fix $\epsilon > 0$

$$\lim P(|X_n - c| \geq \epsilon)$$

$$= \lim_{n \to \infty} [P(X_n \leq c-\epsilon) + P(X_n \geq c+\epsilon)]$$

$$= \lim_{n \to \infty} [P(X_n \leq c-\epsilon)] + \lim_{n \to \infty} [P(X_n \geq c+\epsilon)]$$

$$= \lim_{n \to \infty} F_{X_n}(c-\epsilon) + \lim P(X_n \geq c+\epsilon)$$

$$\leq \underbrace{0}_{\text{from top of page}} + \lim_{n \to \infty} P(X_n > c + \epsilon/2)$$

$c + \epsilon/2$ is subset of $c + \epsilon$ so less strict of inequality

$$= 1 - \lim_{n \to \infty} F_{X_n}(c + \epsilon/2) \qquad \text{Complement rule}$$

$$= 1 - 1$$

$$= 0$$

$$\therefore \lim_{n \to \infty} P(|X_n - c| \geq \epsilon) = 0$$

$$\therefore X_n \xrightarrow{P} c$$

- Thm 5.5.14 (Central Limit Theorem) () part only essential for proof

Let $X_1, \dots, X_n$ be a sequence of iid rvs whose mgfs exist in a neighborhood of 0. (That is, $M_X(t)$ exists for $|t| < h$, for some positive h).

Let $E[X_i] = \mu$ and $Var[X_i] = \sigma^2 > 0$
Define $\bar{X}_n = \frac{1}{n} \Sigma X_i$
Let $G_n(X)$ denote the cdf of $\sqrt{n}(\bar{X}_n - \mu)/\sigma = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X}_n - E[\bar{X}_n]}{Var(\bar{X}_n)}$

Then,

$$\lim_{n \to \infty} G_n(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \, dy$$

$$= \Phi(x) \quad \text{(standard normal)}$$

I.E. $G_n \xrightarrow{P} Z$ where $Z \sim N(0,1)$

Don't say: "The sample is approximately normal when n is large"
Instead: "The dist'n of the sample mean is approx normal when n is large"

- Thm 5.5.17) (Slutsky's Thm.)

If $X_n \xrightarrow{D} X$ in dist'n and
$Y_n \xrightarrow{P} a$, a constant, then

a. $Y_n X_n \xrightarrow{D} aX$
b. $X_n + Y_n \xrightarrow{D} X + a$

Proof: skipped

- Example 5.5.18) Normal Approx w/ estimated Variance

Suppose $X_1, \dots, X_n \overset{iid}{\sim} F$ and
$Var(X_i) \leq \sigma^2 < \infty$

From CLT, we know $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1)$

We know from example 5.5.3 that $S_n^2 \xrightarrow{P} \sigma^2$
Similarly, by theorem 5.5.4 we know that

$S_n = \sqrt{S_n^2} \xrightarrow{P} \sqrt{\sigma^2} = \sigma$ using $h(\cdot) = \sqrt{\cdot}$

Similarly, by Thm 5.5.4: $\frac{\sigma}{S_n} \xrightarrow{P} \frac{\sigma}{\sigma} = 1$ using $h(\cdot) = \frac{\sigma}{\cdot}$

By Slutsky's Theorem:

$$\frac{(\bar{X}_n - \mu)}{S_n/\sqrt{n}} = \underbrace{\frac{\sigma}{S_n}}_{\downarrow P \atop 1} \cdot \underbrace{\frac{(\bar{X} - \mu)}{\sigma/\sqrt{n}}}_{\downarrow D \atop N(0,1)} \quad (\text{mult by } \tfrac{\sigma}{\sigma})$$

$$\xrightarrow{D} 1 \cdot N(0,1) = N(0,1)$$

Consistency of s^2 (using Slutsky's)

Let $X_1, \dots, X_n \overset{iid}{\sim} F$ w/ mean μ and $Var(X_i) = \sigma^2 < \infty$

$$s_n^2 = \frac{1}{n-1} \Sigma (X_i - \bar{X}_n)^2$$

$$= \frac{n}{n-1}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 - \frac{1}{n}\sum_{i=1}(\bar{X}_n-\mu)^2\right]$$

(mult by $\frac{n}{n}$)
(add $\mu - \mu$ in the squared term)

Thm 5.5.4

by WLLN, $\bar{X}_n \xrightarrow{P} \mu \Rightarrow \bar{X}_n - \mu \xrightarrow{P} \mu - \mu = 0$ using $h(\cdot) = \cdot - \mu$

$\therefore (\bar{X}_n - \mu)^2 \xrightarrow{P} 0^2 = 0$ with $h(\cdot) = \cdot^2$

Notation: or $h(x) = x^2$
Prof has used both notations.

Also, by WLLN

$$\frac{1}{n}\sum_{i=1}^{n}\underbrace{(X_i-\mu)^2} \xrightarrow{P} E[(X_i-\mu)^2] = Var(X_i) = \sigma^2$$

$Y_i = X_i - \mu$

$\Rightarrow \bar{Y}_n \xrightarrow{P} E[Y_i]$

$\Rightarrow \bar{Y}_n^2 \xrightarrow{P} E[Y_i^2]$

$\therefore$, by continuous mapping theorem

$$\left.\left[\underbrace{\frac{1}{n}\Sigma(X_i-\mu)^2}_{\downarrow P \; \sigma^2} - \underbrace{\frac{1}{n}\Sigma(\bar{X}_n-\mu)^2}_{\downarrow P \; 0}\right]\right\} ★$$

$$\xrightarrow{P} \sigma^2 + 0 = \sigma^2$$

and, ★ $\xrightarrow{P} \sigma^2$ implies that ★ $\xrightarrow{D} \sigma^2$

Also, $\frac{n}{n-1} \to 1 \Rightarrow \frac{n}{n-1} \xrightarrow{P} 1$

$\therefore$ by Slutsky's

$$\frac{n}{n-1}\left[\frac{1}{n}\Sigma(X_i-\mu)^2 - \frac{1}{n}(\bar{X}_n-\mu)^2\right] \xrightarrow{D} 1 \cdot \sigma^2 = \sigma^2$$

$\therefore s_n^2 \xrightarrow{D} \sigma^2$

Since $\sigma^2 \in \mathbb{R} > 0$,

$\therefore s_n^2 \xrightarrow{P} \sigma^2$ by previous theorem

Thm 5.5.24 (Delta Method)

Let Y_n be a sequence of rvs that satisfies $\sqrt{n}(Y_n - \theta) \xrightarrow{D} N(0, \sigma^2)$

For a given function g and a specific value of θ, Suppose $g'(\theta)$ exists and is not 0. Then

$$\sqrt{n}[g(Y_n) - g(\theta)] \xrightarrow{D} N(0, \sigma^2[g'(\theta)]^2)$$

Example 5.2.25

Suppose that $\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{D} N(0,1)$ w/ $\mu \neq 0$. and $\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{D} N(0,\sigma^2)$

↑ recall, if $Z \sim N(0,1)$, $X = \mu + \sigma Z \sim N(\mu, \sigma^2)$

Then, the delta method says for $g(x) = \frac{1}{x}$, $\Rightarrow g'(x) = -\frac{1}{x^2}$,

① $$\underset{\overset{\uparrow}{g(x)}}{\sqrt{n}\left(\frac{1}{\bar{X}}\right.} - \underset{\overset{\uparrow}{g(\theta)}}{\left.\frac{1}{\mu}\right)} \xrightarrow{D} N\left(0, \sigma^2\left(-\frac{1}{\mu^2}\right)^2\right) = N\left(0, \sigma^2\left(\frac{1}{\mu}\right)^4\right)$$

② Let $g(x) = x^2$, then $g'(x) = 2x$

$\therefore \sqrt{n}(\bar{X}^2 - \mu^2) \xrightarrow{D} (0, \sigma^2 \cdot (2(\mu))^2) = N(0, 4\mu^2\sigma^2)$

What about $1/\bar{X}^2$

Applying the delta method to ① w/ $g(x) = x^2$,

$$\sqrt{n}\left(\left(\frac{1}{\bar{X}_n}\right)^2 - \left(\frac{1}{\mu}\right)^2\right) \xrightarrow{D} N\left(0, \underset{\overset{\uparrow}{\text{from ①}}}{\frac{\sigma^2}{\mu^4}}\underbrace{\left[2\left(\frac{1}{\mu}\right)\right]^2}_{g'(\theta)}\right) = N\left(0, \frac{4\sigma^2}{\mu^6}\right)$$

$Y_n = \frac{1}{\bar{X}_n}$ $\qquad \theta = \frac{1}{\mu}$

Example CLT

CLT says if $X_1, \ldots, X_n \overset{iid}{\sim} F$ w/ mean μ and var $\sigma^2 < \infty$

Then $\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \xrightarrow{D} N(0,1)$

Note: $\sigma \xrightarrow{P} \sigma$. Therefore, by Slutsky's $\underbrace{\sqrt{n}\left(\frac{\bar{X} - \mu}{\sigma}\right) \cdot \sigma}_{\text{exploiting Slutsky's theorem because } \sigma \xrightarrow{P} \sigma} \rightarrow \sigma N(0,1) = N(0,\sigma^2)$

Which takes us to

$$\underbrace{\sqrt{n}(\bar{X}_n - \mu)}_{\text{We want this because it's easy to apply to the delta method!}} \xrightarrow{D} N(0,\sigma^2)$$

Second order delta method

Thm. 5.2.26

Let Y_n be a sequence of rv's such that

$\sqrt{n}(Y_n - \theta) \xrightarrow{D} N(0, \sigma^2)$

For a given function g and a specific value of θ,

suppose $g'(\theta) = 0$, and $g''(\theta)$ exists and is not 0.

Then $n[g(Y_n) - g(\theta)] \xrightarrow{D} \sigma^2 \frac{g''(\theta)}{2} \chi^2_1$ $\quad [\sigma N(0,1)]^2 = \sigma^2 \chi^2_1$

end of ch. 5