Probability and Statistics

(November 9)

Previously done!

i) $X_1, X_2, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

<u>case</u>: μ is unknown, σ^2 is known

known

$\bar{X}$: estimator of μ $\Rightarrow$ $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

unknown

$$Z = \frac{\bar{X} - \mu}{\underbrace{s \; \sigma/\sqrt{n}}} \sim N(0,1)$$

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

estimate σ^2.

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

$S = +\sqrt{S^2}$ estimate of σ.

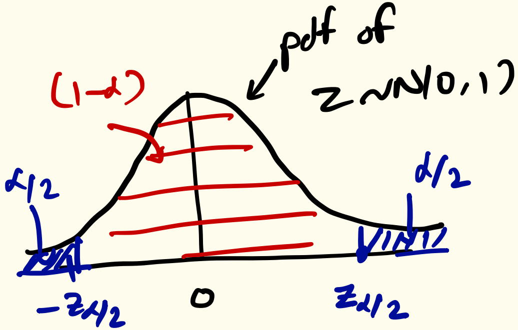
Case (i) $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

$\alpha = 0.1$, $z_{0.05} = 1.65$

μ: unknown

σ^2: known

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



$P(Z > z_{\alpha/2}) = \alpha/2$

$\alpha \in [0,1]$

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1-\alpha$$

$$\Rightarrow P\left(-z_{\alpha/2} < \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(\underbrace{\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}} < \mu < \underbrace{\bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}}\right) = \underbrace{1-\alpha}$$

$L = \bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

$U = \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

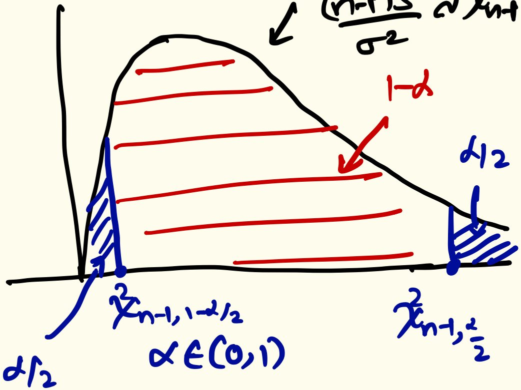
$[L, U]$

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

Estimate : σ^2

Consider : S^2 unbiased estimator of σ^2

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



$$P\left(\chi^2_{n-1,1-\alpha/2} \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{n-1,\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(\frac{1}{\chi^2_{n-1,1-\alpha/2}} \geq \frac{\sigma^2}{(n-1)S^2} \geq \frac{1}{\chi^2_{n-1,\alpha/2}}\right) = 1-\alpha$$

$$\Rightarrow P\left(\underbrace{\frac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}}}_{L} \le \sigma^2 \le \underbrace{\frac{(n-1)S^2}{\chi^2_{n-1,1-\alpha/2}}}_{U}\right) = \underbrace{1-\alpha}$$

case(iii) $X_1, \cdots, X_n$ iid $N(\mu, \sigma^2)$

Here: μ, σ^2 both unknown.

Estimate: μ

$\bar{X}$ is an unbiased estimator of μ.

since σ^2 is also unknown

consider $\frac{\bar{X} - \mu}{S/\sqrt{n}}$ (Replaced σ in Z by $S = \sqrt{S^2}$ estimate of σ)

Q: what is the density of $\frac{\bar{X} - \mu}{S/\sqrt{n}}$??

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} = \frac{\frac{\bar{X} - \mu}{\sigma/\sqrt{n}}}{S/\sigma} = \frac{\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}}{\sqrt{S^2/\sigma^2}} = \frac{Z_1}{\sqrt{Z_2/n-1}}$$

$Z_1 \sim N(0,1)$

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

$$Z_2 = \frac{(n-1)S^2}{\sigma^2}$$

$$Z_2/n-1 = \frac{S^2}{\sigma^2}$$

$$\frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{Z_1}{\sqrt{Z_2/n-1}}$$

where $Z_1 \sim N(0,1)$

$Z_2 \sim \chi^2_{n-1}$

and Z_1 and Z_2 are independent.

$$\boxed{\frac{N(0,1)}{\sqrt{\chi^2_p/p}}}$$

and numerator & denominator are indep.

Aim: Let $U \sim N(0,1)$ and let $V \sim \chi^2_p$ and U & V are independent.

Compute the density of $\dfrac{U}{\sqrt{V/p}}$

$$f_U(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \quad ; \quad -\infty < u < \infty$$

$$f_V(v) = \frac{(1/2)^{p/2}}{\Gamma(p/2)} e^{-v/2} \, v^{p/2 - 1} \quad ; \quad 0 < v < \infty$$

$$= 0 \quad \text{o.w.}$$

Let $f(u,v)$ denote the joint density of U & V.

$$f(u,v) = f_U(u) f_V(v)$$

← since U & V are indep.

$$f(u,v) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2} \frac{1}{\Gamma\left(\frac{p}{2}\right) 2^{p/2}} e^{-v/2} v^{p/2 - 1}$$

<u>transformation</u>: $t = \dfrac{u}{\sqrt{v/p}}$; $w = v$

$$\boxed{\begin{aligned} &\Rightarrow u = t\sqrt{v/p} = t\sqrt{w/p} = t\left(\frac{w}{p}\right)^{1/2} \\ &\Rightarrow v = w \end{aligned}}$$

$-\infty < t < \infty$

$0 < w < \infty$

$$J = \det \begin{bmatrix} \frac{\partial u}{\partial t} & \frac{\partial u}{\partial w} \\ \frac{\partial v}{\partial t} & \frac{\partial v}{\partial w} \end{bmatrix} = \det \begin{bmatrix} \left(\frac{w}{p}\right)^{1/2} & * \\ 0 & 1 \end{bmatrix}$$

$$J = \left(\frac{w}{p}\right)^{1/2}$$

$$f(t,w) = f\left(t\left(\frac{w}{p}\right)^{1/2}, w\right) \times \left(\frac{w}{p}\right)^{1/2}$$

$$f(t,w) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2\frac{w}{p}} \times \frac{1}{\Gamma\left(\frac{p}{2}\right)2^{p/2}} e^{-w/2} w^{p/2-1} \times \left(\frac{w}{p}\right)^{1/2}$$

To compute the density $t = \dfrac{U}{\sqrt{V/p}}$

$$\Rightarrow f(t) = \int_0^\infty f(t,w)\,dw$$

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

$$f(t) = \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma\left(\frac{p}{2}\right)2^{p/2}} \frac{1}{p^{1/2}} \frac{\Gamma\left(\frac{p+1}{2}\right)}{\left(\frac{1}{2}\right)^{\frac{p+1}{2}}\left(\frac{t^2}{p}+1\right)^{\frac{p+1}{2}}}$$

$$\boxed{f(t) = \frac{\Gamma\left(\frac{p+1}{2}\right)}{\Gamma\left(\frac{p}{2}\right)} \frac{1}{\sqrt{p\pi}} \frac{1}{\left(1+\frac{t^2}{p}\right)^{\frac{p+1}{2}}} \quad ; \quad -\infty < t < \infty}$$

pdf of $\frac{U}{\sqrt{V/p}} \sim t_p$ — t density with p degrees of freedom

when $p=1$, Cauchy density

i) Does <u>NOT</u> have MGF

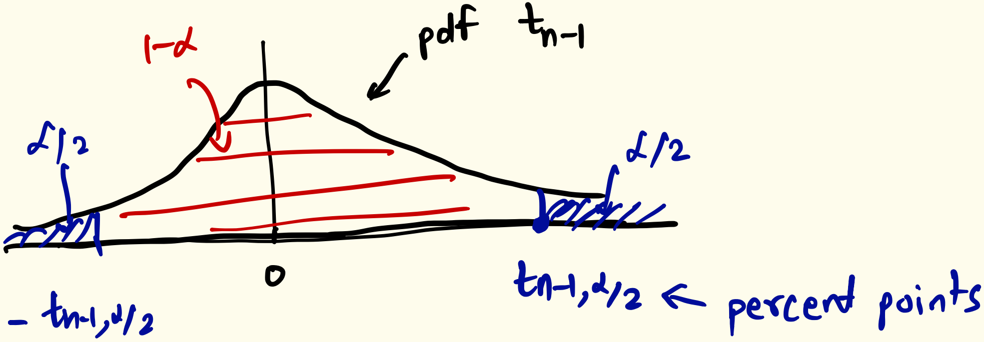
ii) only $p-1$ moments exist for t_p

iii) $E(t_p) = 0$ for $p>1$

$var(t_p) = \frac{p}{p-2}$ for $p>2$

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

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



$\alpha \in (0,1)$

$$P\left(-t_{n-1,\alpha/2} < \frac{\bar{X} - \mu}{S/\sqrt{n}} < t_{n-1,\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\Big(\underbrace{\bar{X} - \frac{S}{\sqrt{n}}\, t_{n-1,\alpha/2}}_{L} < \mu < \underbrace{\bar{X} + \frac{S}{\sqrt{n}}\, t_{n-1,\alpha/2}}_{U}\Big) = 1-\alpha$$

Comparing variances of two distributions/population

$X_1, X_2, \dots, X_n$ is a random sample from $N(\mu_1, \sigma_1^2)$ drawn independently from the random sample $Y_1, Y_2, \dots, Y_m$ from $N(\mu_2, \sigma_2^2)$

Compare σ_1^2 and σ_2^2

$$\frac{S_1^2}{S_2^2}$$

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n-1, m-1}$$

Consider $$\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} = \frac{\frac{(n-1)s_1^2}{\sigma_1^2} \cdot \frac{1}{n-1}}{\frac{(m-1)s_2^2}{\sigma_2^2} \cdot \frac{1}{m-1}} = \frac{U/n-1}{V/m-1}$$

where $U \sim \chi^2_{n-1}$, $V \sim \chi^2_{m-1}$

and further U & V are independent.

In general, let $U \sim \chi^2_p$ and $V \sim \chi^2_q$ and U & V are independent.

Compute the density of $\dfrac{U/p}{V/q}$.

(This density is called as F density with p, q degrees of freedom)

The joint density of U & V

$$f(u,v) = \frac{1}{\Gamma(p/2)\, 2^{p/2}} e^{-u/2} u^{p/2 - 1} \times \frac{1}{\Gamma(\frac{q}{2}) 2^{q/2}} e^{-v/2} v^{q/2 - 1}$$

$0 < u < \infty, \; 0 < v < \infty$

(Since U & V are indep.)

<u>Substitution:</u> $x = \frac{u/p}{v/q}$, $y = v/q$ $\quad$ $0 < x < \infty$, $0 < y < \infty$

$\Rightarrow u = p \cdot x \cdot v/q \Rightarrow u = pxy$

& $v = qy$

$$J = \det \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} = \det \begin{bmatrix} py & px \\ 0 & q \end{bmatrix} = pqy$$

$$f(x,y) = f(pxy,\ qy) \times pqy$$

$$\Rightarrow f(x) = \int_0^\infty \left(\frac{1}{\Gamma(p/2)\, 2^{p/2}} \cdot \frac{1}{\Gamma(\frac{q}{2})\, 2^{q/2}}\, e^{-\frac{pxy}{2}}\, e^{-\frac{qy}{2}}\, (pxy)^{\frac{p}{2}-1} (qy)^{\frac{q}{2}-1}\, pqy \right) dy$$

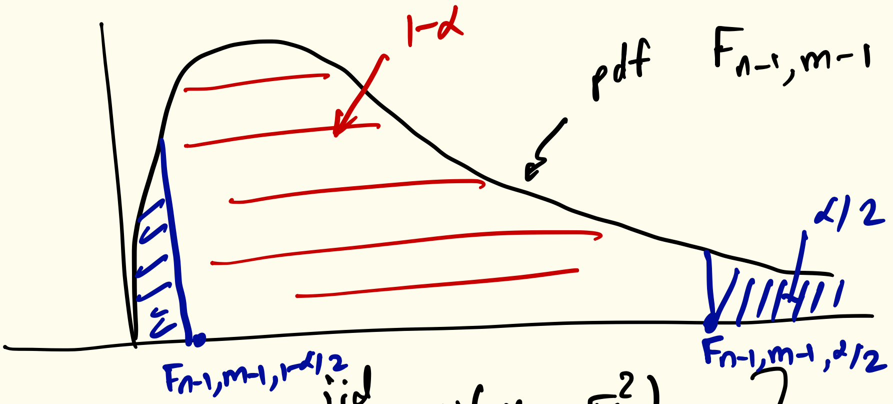
$$= \frac{1}{\Gamma(p/2)\Gamma(q/2)\, 2^{(p+q)/2}} (px)^{p/2-1}\, q^{(q/2-1)}\, pq \times \int_0^\infty e^{-\frac{1}{2}(px+q)y}\, y^{\left(\frac{p+q}{2}\right)-1}\, dy$$

$$f(x) = \frac{\Gamma\left(\frac{p+q}{2}\right)}{\boxed{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)}} \times \frac{1}{\cancel{2^{\frac{p+q}{2}}}} \times p^{p/2} \, q^{q/2} \, x^{p/2 - 1}$$

$$\times \frac{1}{\dfrac{(px+q)^{\frac{p+q}{2}}}{\cancel{2^{\left(\frac{p+q}{2}\right)}}}}$$

$$\boxed{\begin{aligned} f(x) &= \frac{\Gamma\left(\frac{p+q}{2}\right)}{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)} \left(\frac{p}{q}\right)^{p/2} \frac{x^{p/2-1}}{\left(1+\frac{p}{q}x\right)^{\frac{p+q}{2}}} \quad ; \; 0 < x < \infty \\ &= 0 \qquad\qquad \text{o.w.} \end{aligned}}$$

$$X \sim F_{p,q}$$



$$\left.\begin{array}{l} X_1, \ldots, X_n \overset{iid}{\sim} N(\mu_1, \sigma_1^2) \\ Y_1, \ldots, Y_m \overset{iid}{\sim} N(\mu_2, \sigma_2^2) \end{array}\right\} \text{ Independent}$$

$$P\left(F_{n-1,m-1,1-\alpha/2} \le \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \le F_{n-1,m-1,\alpha/2}\right) = 1-\alpha$$

Observe:

$U \sim N(0,1)$; $V \sim \chi^2_p$; U & V independent.

i) $U/\sqrt{V/p} \sim t_p$

ii) $U^2/V/p$ $\qquad$ If $U \sim N(0,1)$

$=$ $\qquad$ $U^2 \sim \chi^2_1$

$U^2/1 \,/\, V/p \sim F_{1,p}$

(square of a t_p random variable is $F_{1,p}$).

Summary:

$X_1, X_2, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

i) μ unknown, σ^2 known

$\bar{X} \sim N(\mu, \sigma^2/n)$ $\qquad$ $\dfrac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$

ii) s^2 is an unbiased estimator of σ^2.

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

iii) $\dfrac{\bar{X} - \mu}{s/\sqrt{n}} \sim t_{n-1}$

$X_1, \dots, X_n \overset{iid}{\sim} N(\mu_1, \sigma_1^2)$
$Y_1, \dots, Y_m \overset{iid}{\sim} N(\mu_2, \sigma_2^2)$ } independ

$$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F_{n-1,m-1}$$

$X \sim N(\mu, \sigma^2)$

$\Rightarrow \frac{X-\mu}{\sigma} \sim N(0,1)$

In general:

i) let $Z_1, Z_2, \dots, Z_p \overset{iid}{\sim} N(0,1)$

$$Z_1^2 + Z_2^2 + \dots + Z_p^2 \sim \chi^2_p = \text{gamma}\left(\frac{p}{2}, \frac{1}{2}\right)$$

(sum of squares of p independent standard normal r.v.s. follow χ^2_p).

ii) $U \sim N(0,1)$; $V \sim \chi^2_p$; U & V are indep.

$$U \Big/ \sqrt{V/p} \sim t_p$$

iii) $U \sim \chi^2_p$, $V \sim \chi^2_q$, U & V are indep.

$$\frac{U/p}{V/q} \sim F_{p,q}$$

Assignment 8 (P5)

$X_1, \cdots, X_n, X_{n+1}$ are iid $N(\mu, \sigma^2)$

$\bar{X}$ and s^2 : sample mean & sample variance based $X_1, \cdots, X_n$

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}, \qquad s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Q: Find the distribution of $\sqrt{\frac{n}{n+1}} \left(\frac{X_{n+1} - \bar{X}}{s} \right)$

X_{n+1} is indep. of $\bar{X}$ and s^2.

$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$, $X_{n+1} \sim N(\mu, \sigma^2)$

$$x_{n+1} - \bar{x} \sim N\left(0, \underbrace{\frac{\sigma^2}{n} + \sigma^2}\right)$$

$$\Rightarrow \quad x_{n+1} - \bar{x} \sim N\left(0, \sigma^2\left(\frac{n+1}{n}\right)\right)$$

$$\Rightarrow \quad \frac{x_{n+1} - \bar{x}}{\sigma\sqrt{\frac{n+1}{n}}} \sim N(0,1)$$

Note that $x_{n+1} - \bar{x}$ is indep. of s^2

$$\frac{x_{n+1} - \bar{x}}{s\sqrt{\frac{n+1}{n}}} = \boxed{\frac{x_{n+1} - \bar{x}/\sigma}{\sqrt{s^2/\sigma^2}} \cdot \sqrt{\frac{n}{n+1}}} \sim t_{n-1}$$

$$\Rightarrow \frac{N(0,1)}{\sqrt{\chi^2_{n-1}/n-1}}$$

since $(n-1)s^2/\sigma^2 \sim \chi^2_{n-1}$

Ex: Let $X_1, \dots, X_n$ be random sample from $N(\mu, \sigma^2)$

for $1 < k < n$,

$$U = \frac{1}{k} \sum_{i=1}^{k} X_i \quad ; \quad V = \frac{1}{n-k} \sum_{i=k+1}^{n} X_i$$

$$S^2 = \frac{1}{k-1} \sum_{i=1}^{k} (X_i - U)^2 \quad ; \quad T^2 = \frac{1}{n-k-1} \sum_{i=k+1}^{n} (X_i - V)^2$$

$$\underbrace{\boxed{X_1, \dots, X_k}}_{k \text{ indep. r.v.s.}} \qquad \underbrace{\boxed{X_{k+1}, \dots, X_n}}_{n-k \text{ indep r.v.s.}}$$

$$U \sim N\left(\mu, \frac{\sigma^2}{k}\right) \quad ; \quad V \sim N\left(\mu, \frac{\sigma^2}{n-k}\right)$$

$$\frac{(k-1)s^2}{\sigma^2} \sim \chi^2_{k-1} \quad ; \quad \frac{(n-k-1)T^2}{\sigma^2} \sim \chi^2_{n-k-1}$$

i) density of $\frac{U+V}{2}$

U & V are indep.

$U \sim N(\mu, \sigma^2/k)$, $\quad V \sim N(\mu, \sigma^2/n-k)$

$$\frac{U+V}{2} \sim N\left(\mu, \frac{1}{4}\left(\frac{\sigma^2}{k} + \frac{\sigma^2}{n-k}\right)\right)$$

ii) Density $W_2 = \dfrac{(k-1)S^2 + (n-k-1)T^2}{\sigma^2}$

$$= \frac{(k-1)S^2}{\sigma^2} + \frac{(n-k-1)T^2}{\sigma^2}$$

$\sim \chi^2_{k-1}$ $\sim \chi^2_{n-k-1}$

We know S^2 & T^2 are indep.

$$W_2 \sim \chi^2_{(k-1)+(n-k-1)}$$

$$\Rightarrow W_2 \sim \chi^2_{n-2}$$