Lec 18 Math 621 11/20/19

$X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2) \Rightarrow Z_1 = \frac{X_1 - \mu}{\sigma}, \ldots Z_n = \frac{X_n - \mu}{\sigma} \overset{iid}{\sim} N(0,1)$

$$\underbrace{\vec{Z}^T \vec{Z}}_{\sim \chi^2_n} = \sum Z_i^2 = \sum \left(\frac{X_i - \mu}{\sigma}\right)^2 = \frac{\sum (X_i - \mu)^2}{\sigma^2} = \frac{(n-1) S_n^2}{\sigma^2} + \underbrace{\frac{n(\bar{X} - \mu)^2}{\sigma^2}}_{\sim \chi^2_1} \sim \chi^2_n$$

Recall $\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \Rightarrow \underbrace{\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)}_{\text{justification for the "Z test for one-sample means"}} \Rightarrow \left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right)^2 \sim \chi^2_1$

$$\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right)^2 = \frac{(\bar{X} - \mu)^2}{\frac{\sigma^2}{n}} = \frac{n(\bar{X} - \mu)^2}{\sigma^2}$$

We know if $U_1 \sim \chi^2_{k_1}$ indep of $U_2 \sim \chi^2_{k_2} \Rightarrow U_1 + U_2 \sim \chi^2_{k_1 + k_2}$

Conjecture: ① $\frac{(n-1) S_n^2}{\sigma^2} \sim \chi^2_{n-1}$ and ② indep of $\frac{n(\bar{X} - \mu)^2}{\sigma^2}$

i.e. S_n^2 and $\bar{X}_n$ are indep. since n, σ^2 are just constants

$$\frac{(n-1)S^2}{\sigma^2} + \underbrace{\frac{n(\bar{X}-\mu)^2}{\sigma^2}}^{\chi^2_1} = \frac{\sum(X_i-\mu)^2}{\sigma^2} \sim \chi^2_n$$

Conjecture: $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ and indep. of $\frac{n(\bar{X}-\mu)^2}{\sigma^2}$

Consider...

quadratic form

$$\vec{Z}^T\vec{Z} \sim \chi^2_n \Rightarrow \vec{Z}^T I_n \vec{Z} \sim \chi^2_n$$

Consider

$$\vec{Z}^T \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \vdots \\ \vdots & & & \\ 0 & \cdots & & 0 \end{bmatrix}}_{B_1} \vec{Z} = Z_1^2 \sim \chi^2_1$$

$$\vec{Z}^T \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & & \\ 0 & \cdots & & 0 \end{bmatrix}}_{B_2} \vec{Z} = Z_2^2 \sim \chi^2_1$$

all indep. since $Z_1, \ldots, Z_n \overset{iid}{\sim}$

$$\vdots$$

$$\vec{Z}^T \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}}_{B_n} \vec{Z} = Z_n^2 \sim \chi^2_1$$

$$\Rightarrow \vec{Z}^T B_1 \vec{Z} + \vec{Z}^T B_2 \vec{Z} + \cdots + \vec{Z}^T B_n \vec{Z} \sim \chi^2_n$$

Note $\vec{Z}^T \underbrace{(B_1 + B_2 + \cdots + B_n)}_{I} \vec{Z} \sim \chi^2_n$

$$\text{rank}[B_1] = 1,\ \text{rank}(B_2) = 1, \ldots, \text{rank}(B_n) \Rightarrow \sum_{i=1}^{n} \text{rank}(B_i) = n$$

Cochran's Thm

If $B_1 + \cdots + B_n = I_n$ and $\sum \text{rank}(B_i) = n \Rightarrow$ (a) $\vec{Z}^T B_i \vec{Z} \sim \chi^2_{\text{rank}[B_i]}$

(b) $\vec{Z}^T B_i \vec{Z}$ indep of $\vec{Z}^T B_j \vec{Z}$ $\forall i \neq j$

Proof on HW for MA students

Consider: $\vec{z}^T\vec{z} = \sum z_i^2 = \sum((z_i-\bar{z})+\bar{z})^2$

$= \sum (z_i-\bar{z})^2 + 2(z_i-\bar{z})\bar{z} + \bar{z}^2$

$= \sum(z_i-\bar{z})^2 + 2\sum z_i\bar{z} - \bar{z}^2 + \sum\bar{z}^2$

$= \sum(z_i-\bar{z})^2 + 2\cancelto{0}{(n\bar{z}^2 - n\bar{z}^2)} + n\bar{z}^2$

Let $\vec{1}_n := \begin{pmatrix}1\\ \vdots \\ 1\end{pmatrix}$ $\quad \bar{z} = \frac{z_1+\dots z_n}{n} = \frac{1}{n}\vec{1}^T\vec{z} = \frac{1}{n}\vec{z}^T\vec{1}_n$

$n\bar{z}^2 = n\bar{z}\bar{z} = \cancel{n}\left(\frac{1}{\cancel{n}}\vec{z}^T\vec{1}_n\right)\left(\frac{1}{n}\vec{1}_n^T\vec{z}\right) = \vec{z}^T\left(\frac{1}{n}\vec{1}_n\vec{1}_n^T\right)\vec{z} = \vec{z}^T\left(\frac{1}{n}J_n\right)\vec{z}$

Let $J_n := \begin{bmatrix}1 & \cdots & 1\\ \vdots & & \vdots \\ 1 & \cdots & 1\end{bmatrix} = \vec{1}\vec{1}_n$

$\sum(z_i-\bar{z})^2 = \sum z_i^2 - 2z_i\bar{z} + \bar{z} = \sum z_i^2 - 2n\bar{z}^2 + n\bar{z}^2 = \sum z_i^2 - n\bar{z}^2$

$= \vec{z}^T\vec{z} - \vec{z}^T\left(\frac{1}{n}J_n\right)\vec{z} = \vec{z}^T\left(\vec{z} - \frac{1}{n}J_n\vec{z}\right) = \vec{z}^T\left(I_n - \frac{1}{n}J_n\right)\vec{z}$

$\vec{z}^T\vec{z} = \vec{z}^T\underbrace{\left(I_n - \frac{1}{n}J_n\right)}_{B_1}\vec{z} + \vec{z}^T\underbrace{\left(\frac{1}{n}J_n\right)}_{B_2}\vec{z}$

$B_1 = \begin{bmatrix}1-\frac{1}{n} & -\frac{1}{n} & \cdots & -\frac{1}{n}\\ -\frac{1}{n} & 1-\frac{1}{n} & & \vdots \\ -\frac{1}{n} & & \ddots & 1-\frac{1}{n}\end{bmatrix}$ symmetric

$B_1 + B_2 = I_n$ Do the ranks add?

$B_2 = \begin{bmatrix}\frac{1}{n} & \frac{1}{n}\\ \frac{1}{n} & \frac{1}{n}\end{bmatrix}$ symmetric

Thm: if A is symmetric and idempotent (i.e. $AA=A$) then $\text{rank}(A) = tr[A]$

$B_2B_2 = \frac{1}{n}J_n\frac{1}{n}J_n = \frac{1}{n^2}J_nJ_n = \frac{1}{n^2}nJ_n = \frac{1}{n}J_n$

$B_1B_1 = \left(I_n - \frac{1}{n}J_n\right)\left(I_n - \frac{1}{n}J_n\right) = I_nI_n - \frac{1}{n}I_nJ_n - \frac{1}{n}J_nI_n + \frac{1}{n^2}J_nJ_n$

$= I - 2\left(\frac{1}{n}J_n\right) + \frac{1}{n}J_n = I - \frac{1}{n}J_n$

$\text{rank}(B_1) = n\left(1-\frac{1}{n}\right) = n-1$

$+$

$\text{rank}(B_2) = n\frac{1}{n} = 1$

n

$\Rightarrow \sum(Z_i - \bar{Z})^2 \sim \chi^2_{n-1}$ indep of $n\bar{Z}^2 \sim \chi^2_1$

Note: $\bar{Z} = \dfrac{Z_1 + \dots + Z_n}{n} = \dfrac{\frac{X_1-\mu}{\sigma} + \dots + \frac{X_n-\mu}{\sigma}}{n} = \dfrac{X_1 + \dots + X_n - n\mu}{n\sigma}$

$= \dfrac{n\left(\frac{X_1+\dots+X_n}{n} - \mu\right)}{n\sigma} = \dfrac{\bar{X}-\mu}{\sigma}$

$\Rightarrow \sum\left(Z_i - \dfrac{\bar{X}-\mu}{\sigma}\right)^2 = \sum\left(\dfrac{X_i-\mu}{\sigma} - \dfrac{\bar{X}-\mu}{\sigma}\right)^2 = \dfrac{\sum(X_i-\bar{X})^2}{\sigma^2}$

$= \dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ which proves our conjecture.

and justifies the χ^2 test for one sample variances

Aside

$\dfrac{S^2_{n_1}/\sigma^2_1}{S^2_{n_2}/\sigma^2_2} \sim F_{n_1,n_2}$

justifies 2-sample variance test

e.g. $H_0: \sigma^2_1 = \sigma^2_2$

$\dfrac{S^2_1}{S^2_2} \sim F_{n_1,n_2}$

thus $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ and $\dfrac{n(\bar{X}-\mu)^2}{\sigma^2} \sim \chi^2_1$ are indep.

$\Rightarrow \dfrac{\sqrt{n-1}}{\sigma} S \sim \chi_{n-1}$ "Chi" and $S^2 \sim \text{Gamma}\left(\frac{n-1}{2}, \frac{n-1}{2\sigma^2}\right)$ HW

$\Rightarrow$ Since n, μ, σ^2 are constants $\Rightarrow S^2$ and $\bar{X}$ are independent.

First proved by Fisher in 1925. Geary in 1936 proved that $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ is the only distr. that allows for this.

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

$\dfrac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$ this allows for the one-sample Z-test in the basic stat class

$\dfrac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} \sim ? = \dfrac{\bar{X}-\mu}{\sqrt{\frac{S^2}{n}}} = \dfrac{\bar{X}-\mu}{\frac{1}{\sqrt{n}}\sqrt{S^2}} = \dfrac{\bar{X}-\mu}{\frac{1}{\sqrt{n}}\sqrt{\frac{\sigma^2}{n-1}\frac{n-1}{\sigma^2}S^2}} = \dfrac{\left.\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}\right\} \sim N(0,1) \text{ a independent}}{\sqrt{\dfrac{\left.\frac{(n-1)S^2}{\sigma^2}\right\} \sim \chi^2_{n-1}}{n-1}}}$

σ unknown so plug in S. (Gosset's idea!)

$= T_{n-1}$ One sample T-test !!

Multivariate Normal

let $\vec{Z} := \begin{bmatrix} Z_1 \\ \vdots \\ Z_n \end{bmatrix}$ where $Z_1, \ldots, Z_n \overset{iid}{\sim} N(0,1)$

$$E[\vec{Z}] = \vec{0}_n \quad , \quad Var[\vec{Z}] = I_n \quad , \quad \vec{Z} \sim f_{\vec{Z}}(\vec{z}) = ?$$

Since we have iid...

$$f_{Z_1, \ldots, Z_n}(z_1, \ldots, z_n) = \prod_{i=1}^{n} f(z_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_i^2}{2}} = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2} \vec{z}^T \vec{z}}$$

multivariate normal, expectation, variance

$$= N_n(\vec{0}_n, I_n)$$

dimension

Consider $\vec{\mu} \in \mathbb{R}^n$

let $\vec{X} = \vec{Z} + \vec{\mu} = \begin{bmatrix} Z_1 + \mu_1 \\ Z_2 + \mu_2 \\ \vdots \\ Z_n + \mu_n \end{bmatrix} \begin{matrix} \sim N(\mu_1, 1) \\ \sim N(\mu_2, 1) \\ \vdots \\ \sim N(\mu_n, 1) \end{matrix} \Rightarrow \vec{X} \sim N_n(\vec{\mu}, I_n)$

$E[\vec{X}] = E[\vec{Z}] + \vec{\mu} = \vec{0} + \vec{\mu} = \vec{\mu}$

why? still independent

let $A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & 0 \cdots & 0 \\ \vdots & & & & \\ 1 & 1 & \cdots & \cdots & 1 \end{bmatrix}}_{n \times n}$ let $\vec{X} = A\vec{z} = \begin{bmatrix} Z_1 \\ Z_1 + Z_2 \\ Z_1 + Z_2 + Z_3 \\ \vdots \\ Z_1 + \cdots + Z_{n-1} \\ Z_1 + \cdots + Z_n \end{bmatrix} \begin{matrix} \sim N(0,1) \\ \sim N(0,2) \\ \sim N(0,3) \\ \\ \sim N(0, n-1) \\ \sim N(0, n) \end{matrix}$

Is X_1, X_2 ind? Is Z_1 indep of $Z_1 + Z_2$? NO!

$$Cov[Z_1, Z_1 + Z_2] = \underbrace{Cov(Z_1, Z_1)}_{= Var(Z_1)} + \underbrace{Cov(Z_1, Z_2)}_{0} = 1 \neq 0$$

What is $E(\vec{X}) = E(A\vec{Z})$?

What is $Var(\vec{X}) = Var(A\vec{Z})$?