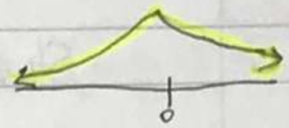
Wednesday 18th November 2020

Lecture 19

$X \sim \text{Cauchy}(0,1) = \frac{1}{\pi}\frac{1}{x^2+1}$

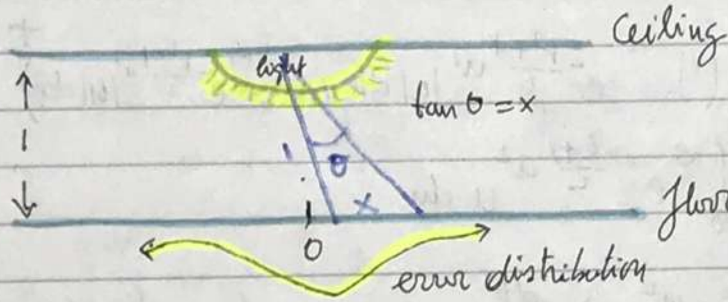
$E[X] = \int_{\mathbb{R}} x \frac{1}{\pi}\frac{1}{x^2+1}\,dx = \infty$ the expectation doesn't exist

$M_X(t) = \int_{\mathbb{R}} e^{tx}\frac{1}{\pi}\frac{1}{x^2+1}\,dx$ does not exist



$\phi_X(t) = \int_{\mathbb{R}} e^{itx}\frac{1}{\pi}\frac{1}{x^2+1}\,dx = \cdots\cdots = e^{-|t|}$, $\phi_X'(t) = -\frac{1}{|t|}e^{-|t|}$, $\phi_X'(0)$ dne

Let's derive the Cauchy distribution like the physicists found it.



$\theta \sim U(-\frac{\pi}{2}, \frac{\pi}{2}) = \frac{1}{\pi}\mathbb{1}_{\theta\in[-\frac{\pi}{2},\frac{\pi}{2}]}$

$x = g(\theta)$

$\theta = g^{-1}(x) = \arctan(x)$

tangent is invertible between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$.

$$f_X(x) = f_\theta(g^{-1}(\theta))\left|\frac{d}{dx}[g^{-1}(x)]\right| = \frac{1}{\pi}\underbrace{\mathbb{1}_{\arctan(x)\in[-\frac{\pi}{2},\frac{\pi}{2}]}}_{x\in\mathbb{R}}\frac{1}{x^2+1} = \text{Cauchy}(0,1)$$

Let $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2) \Rightarrow \frac{X_i - \mu}{\sigma} = Z_i \sim N(0,1)$

$\leftarrow$ Sample Variance

$T_n \sim N(n\mu, n\sigma^2)$, $\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$, $S^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2 \sim f_{S_n^2}(s^2) = ?$

dot product

$\Rightarrow Z_1, \dots, Z_n \overset{iid}{\sim} N(0,1)$ $\vec{Z}^T\vec{Z}$ $\sum_{i=1} Z_i^2 \sim \chi^2_n = \sum_{i=1}^n \left(\frac{X_i-\mu}{\sigma}\right)^2 = \frac{\sum(X_i-\mu)^2}{\sigma^2}$

$(X_i-\mu)^2 = ((X_i - \bar{X}) + (\bar{X} - \mu))^2 = (X_i - \bar{X})^2 + \sum 2(X_i - \bar{X})(\bar{X} - \mu) + \sum(\bar{X} - \mu)^2$

$\sum X_i\bar{X} - \bar{X}^2 - X_i\mu + \bar{X}\mu$

$= \underbrace{\cancel{\bar{X}\sum X_i}}_{n\bar{X}^2} - n\bar{X}^2 - \underbrace{\cancel{\mu\sum X_i}}_{n\bar{X}\mu} + n\bar{X}\mu$

$$= \Sigma (x_i - \bar{x})^2 + n(\bar{x} - \mu)^2$$

$$\Rightarrow \frac{\Sigma (x_i - \mu)^2}{\sigma^2} = \underbrace{\frac{n-1}{\sigma^2} S^2}_{\chi^2_{n-1}} + \underbrace{\left(\frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}\right)^2}_{Z^2 \sim \chi^2_1} \sim \chi^2_n \qquad \frac{n(\bar{X} - \mu)^2}{\sigma^2} = \left(\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma}\right)^2 = \left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}\right)^2$$

Maybe....

$U_1 \sim \chi^2_{k_1}$ indep of $U_2 \sim \chi^2_{k_2}$

$\Rightarrow U_1 + U_2 \sim \chi^2_{k_1 + k_2}$

In order for this "maybe" to be true, we need independence of these two term two terms i.e. we need S^2 and xbar to be independent. We need Cochran's Theorem to prove this.

$\vec{Z}^T \vec{Z} = \vec{Z}^T I \vec{Z} \sim \chi^2_n$ · this scalar is called a "quadratic form"

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

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

rank $[B_i] = 1$

Σ rank $[B_i] = n$

$$\vec{Z}^T \vec{Z} = \vec{Z}^T (B_1 + B_2 + \dots + B_n) \vec{Z} = \underbrace{\vec{Z}^T B_1 \vec{Z}}_{\chi^2_1} + \underbrace{\vec{Z}^T B_2 \vec{Z}}_{\chi^2_1} + \dots + \underbrace{\vec{Z}^T B_n \vec{Z}}_{\chi^2_1} \sim \chi^2_n$$

Conjecture: each of these quadratic forms is independent. ~~Cochran's Theorem~~

Cochran's Thm: If $B_1 + B_2 + \dots + B_K = I$, $K \leq n$ and sum of their ranks is n then you have two powerful results:

a) $\vec{Z}^T B_j \vec{Z} \sim \chi^2_{\text{rank}[B_j]}$ and b) $\vec{Z}^T B_{j_1} \vec{Z}$ is independent of $\vec{Z}^T B_{j_2} \vec{Z}$ $\forall j_1 \neq j_2$

Consider ~~[illegible]~~ $\sum (z_i - \bar{z})^2 = \sum z_i^2 - 2\sum \underbrace{z_i}_{n\bar{z}} \bar{z} + \sum \bar{z}^2 = \sum z_i^2 - 2n\bar{z}^2 + n\bar{z}^2$

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

Let $\vec{1}_n$ = n-dim column vector of all ones $\bar{z} = \frac{1}{n}\vec{1}^T\vec{z} = \frac{1}{n}\vec{z}^T\vec{1}$

$n\bar{z}^2 = n\bar{z}\bar{z} = \cancel{n}\frac{1}{\cancel{n}}\vec{z}^T\vec{1}\frac{1}{n}\vec{1}^T z = \frac{1}{n}\vec{z}\vec{1}\vec{1}^T\vec{z} = \vec{z}^T\underbrace{(\tfrac{1}{n}J_n)}_{B_2}\vec{z}$

Let $J_n = \vec{1}\vec{1}^T$, which is an $n \times n$ matrix of all ones

$\cancel{\vec{z}^T I\vec{z} = \sum z_i^2 + n\bar{z}^2 - n\bar{z}^2}$

$\cancel{= \vec{z}^T\vec{z} + \vec{z}^T(\tfrac{1}{n}J_n)\vec{z} - \vec{z}^T(\tfrac{1}{n}J_n)\vec{z}}$

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

$\cancel{\vec{z}^T I\vec{z} =}$

$\sum (z_i - \bar{z})^2 = \vec{z}^T I\vec{z} - \vec{z}^T(\tfrac{1}{n}J_n)\vec{z} = \vec{z}^T\underbrace{(I - \tfrac{1}{n}J_n)}_{B_1}\vec{z}$

$\vec{z}^T\vec{z} = \sum (z_i - \bar{z})^2 + n\bar{z}^2 = \vec{z}^T B_1\vec{z} + \vec{z} B_1\vec{z}$

I want to use Cochran's thm on the above expression. So I need to make sure $B_1 + B_2 = I$ and $\text{rank}[B_1] + \text{rank}[B_2] = n$

$B_1 + B_2 = (I - \frac{1}{n}J_n) + \frac{1}{n}J_n = I$ ✓

$\text{rank}[B_2] = \text{rank}[\frac{1}{n}J_n] = \text{rank}[J] = 1$

$\therefore \sum \text{rank}[B_j] = 1 + n - 1 = n$

$\text{rank}[B_1] = \text{rank}[I - \frac{1}{n}J]$

Thm from 231 class: if A is symmetric and idempotent (ie $AA = A$) then $\text{rank}[A] = \text{Tr}[A]$ = sum of A's diagonal entries.

$$(I - \tfrac{1}{n} J)^T = I^T - \tfrac{1}{n} J^T = I - \tfrac{1}{n} J \checkmark \qquad J_3 J_3 = \begin{bmatrix} 1&1&1\\1&1&1\\1&1&1 \end{bmatrix}\begin{bmatrix} 1&1&1\\1&1&1\\1&1&1 \end{bmatrix} = \begin{bmatrix} 3&3&3\\3&3&3\\3&3&3 \end{bmatrix} = 3J$$

$$(I - \tfrac{1}{n} J)(I - \tfrac{1}{n} J) = II - \tfrac{1}{n} JI - \tfrac{1}{n} IJ + \tfrac{1}{n^2} JJ = I - \tfrac{2}{n} J + \tfrac{1}{n^{\cancel{2}}} \cancel{n} J$$

$$= I - \tfrac{1}{n} J$$

$$tr[I - \tfrac{1}{n} J] = (1 - \tfrac{1}{n} 1) + (1 - \tfrac{1}{n} 1) + \cdots + (1 - \tfrac{1}{n} 1) = n-1$$

$$= rank[B_1]$$

Since the two conditions of Cochran's Thm are satisfied, we can apply it to get the two results:

$$\Rightarrow \vec{Z}^T B_1 \vec{Z} = \sum (Z_i - \bar{Z})^2 \sim \chi^2_{n-1} \text{ ind of } \vec{Z}^T B_2 \vec{Z} = n \bar{Z}^2 \sim \chi^2_1$$

What does this have to do with our goal? well, it's the same thing:

$$\bar{Z} = \frac{Z_1 + \cdots + Z_n}{n} = \frac{\frac{x_1 - \mu}{\sigma} + \cdots + \frac{x_n - \mu}{\sigma}}{n} = \frac{\sum x_i - n\mu}{\sigma n} = \frac{\bar{X} - \mu}{\sigma}$$

$$\sum (Z_i - \bar{Z})^2 = \sum \left(\frac{x_i - \mu}{\sigma} - \frac{\bar{X} - \mu}{\sigma} \right)^2 = \sum \left(\frac{x_i - \bar{X}}{\sigma} \right)^2 = \frac{1}{\sigma^2} \sum (x_i - \bar{X})^2 = \frac{n-1}{\sigma^2} S^2$$

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

$$\underbrace{\frac{n-1}{\sigma^2} S^2}_{\sim \chi^2_{n-1}} + \underbrace{\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \right)^2}_{\sim \chi^2_1} \sim \chi^2_n$$

independent

Fisher proved this without Cochran's thm in 1925 and Geary proved in 1936 that this decomposition is exclusive to the iid normal r.v model.

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1), \quad \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim \textbf{?} \quad \text{Not } N(0,1)$$