→ **Important Statistics & Sampling Distributions for a normal (m, σ) population.**

① $U = \dfrac{\bar{X} - m}{\sigma/\sqrt{n}} \sim \text{Normal}(0,1)$ (using Central Limit Thm).

$X_1, X_2, \ldots, X_n$ all have ~~[illegible]~~ identical independent normal (m, σ) distributions (Actually also, it is normal$(m, \frac{\sigma}{\sqrt{n}})$. proof in PS_1 Q3).

② $\chi^2 = \dfrac{nS^2}{\sigma^2} \sim \chi^2(n-1)$. How?.

Proof ⇒ $S^2 = \dfrac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$

$$\Rightarrow nS^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2 \Rightarrow \frac{nS^2}{\sigma^2} = \sum_{i=1}^{n} \left(\frac{X_i - \bar{X}}{\sigma}\right)^2$$

$$\cancel{= \sum_{i=1}^{n} \left(\frac{X_i}{\sigma}\right)^2 + n \cdot \left(\frac{\bar{X}}{\sigma}\right)^2 - 2 \cdot \frac{\bar{X}}{\sigma^2} \cdot \underbrace{\sum_{i=1}^{n} X_i}_{n\bar{X}}}$$

$$\cancel{= \sum_{i=1}^{n} \left(\frac{X_i}{\sigma}\right)^2 - n\left(\frac{\bar{X}}{\sigma}\right)^2}$$

$$= \sum_{i=1}^{n} \left(\frac{(X_i - m) - (\bar{X} - m)}{\sigma}\right)^2$$

$$= \sum_{i=1}^{n}\left(\frac{X_i-m}{\sigma}\right)^2 + n\cdot\left(\frac{\bar{X}-m}{\sigma}\right)^2 \quad \text{--- (1)}$$

$$-2\cdot\sum_{i=1}^{n}\left(\frac{X_i-m}{\sigma}\right)\cdot\left(\frac{\bar{X}-m}{\sigma}\right)$$

$$-2\cdot\left(\frac{\bar{X}-m}{\sigma}\right)\cdot\left[\sum_{i=1}^{n}\frac{X_i}{\sigma} - \frac{nm}{\sigma}\right]$$

$$= -2\left(\frac{\bar{X}-m}{\sigma}\right)\cdot\left[\frac{n\bar{X}}{\sigma} - \frac{nm}{\sigma}\right]$$

$$= -2\cdot\left(\frac{\bar{X}-m}{\sigma}\right)^2\cdot n.$$

$$\therefore (1) = \sum_{i=1}^{n}\left(\frac{X_i-m}{\sigma}\right)^2 - n\left(\frac{\bar{X}-m}{\sigma}\right)^2.$$

$$= \sum_{i=1}^{n}\left(\frac{X_i-m}{\sigma}\right)^2 - \left(\frac{\bar{X}-m}{\sigma/\sqrt{n}}\right)^2.$$

Now $\frac{X_i-m}{\sigma} \sim$ Normal $(0,1)$ Normal $(0,1)$.

We know for S.N. Variates $X_1^2+X_2^2+\dots+X_n^2 \sim \chi^2(n)$ variate.

$\therefore (1) \Rightarrow \chi^2(n) - \chi^2(1) \Rightarrow \boxed{\frac{nS^2}{\sigma^2} \sim \chi^2(n-1)}$ Hence Proved.

(3) $t = \frac{\bar{X}-m}{s/\sqrt{n}} \sim t(n-1)$ where $s^2 = \frac{n}{n-1}S^2$.

<u>Proof</u> :- $\frac{\bar{X}-m}{\sigma/\sqrt{n}} \sim$ Normal $(0,1)$.

$\frac{nS^2}{\sigma^2} \Rightarrow \chi^2(n-1)$ variate.

Now $\frac{nS^2}{\sigma^2} \sim \chi^2(n-1)$

We know that $\frac{X}{\sqrt{\frac{Y}{n}}} \sim t(n)$ if $X \sim$ Normal $(0,1)$
$Y \sim \chi^2(n)$.

$$\therefore \frac{\left(\dfrac{\bar{X}-m}{\sigma/\sqrt{n}}\right)}{\sqrt{\dfrac{nS^2}{\sigma^2(n-1)}}} \sim t(n-1). \quad \Rightarrow \quad \frac{\dfrac{\bar{X}-m}{\cancel{\sigma}/\sqrt{n}}}{\dfrac{\sqrt{n}.S}{\cancel{\sigma}.\sqrt{n-1}}}$$

It is given that $s^2 = \frac{n}{n-1} . S^2 \Rightarrow s = \frac{\sqrt{n}.S}{\sqrt{n-1}}$.

$$\therefore \frac{\bar{X}-m}{s} . \sqrt{n} \sim t(n-1) \Rightarrow \boxed{\frac{\bar{X}-m}{s/\sqrt{n}} \sim t(n-1)}$$

Hence Proved.

Problem Set 1 (Stats)

Q1. $P(X=0) = P(X=1) = 1/2$.

Random sample of size 4 is drawn from population of x.

To show $\Rightarrow$ $t = x_1 + x_2 + x_3 + x_4$ is $\sim$ Binomial$\left(4, \frac{1}{2}\right)$.

$t =$ ~~Average~~ Sum of x_1, x_2, x_3, x_4

$t = 0$ when all are 0.

$P(t=0) = P(x_1=0) \cdot P(x_2=0) \cdot P(x_3=0) \cdot P(x_4=0)$

$= \boxed{\frac{1}{16}} = {}^4C_0 \cdot \left(\frac{1}{2}\right)^4$

$P(t=1) = P(\text{one out of 4 being 1})$

$= {}^4C_1 \cdot \left(\frac{1}{2}\right)^4 \Rightarrow$ which one is 1.

IIIly $P(t=2) = {}^4C_2 \cdot \left(\frac{1}{2}\right)^4$. $\quad P(t=3) = {}^4C_3 \cdot \left(\frac{1}{2}\right)^4$.

$P(t=4) = {}^4C_4 \cdot \left(\frac{1}{2}\right)^4$.

$\therefore$ Clearly $t \sim \text{Binomial}(4, 1/2)$.

~~Q2. Show that Sample Mean is asymptotically normal.~~
~~Done already. (Proof for $\frac{\bar{X}-m}{\sigma/\sqrt{n}} \sim \text{Normal}(0,1)$).~~

Q2. Let $(x_1, x_2, \ldots, x_n)$ is the sample.

$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$ $\Rightarrow$ $\bar{X}$ = Sampling Distribution $= \frac{X_1 + X_2 + \cdots + X_n}{n}$.

where each X_i is identical.

By Central Limit Thm.

$\left(\frac{\bar{X} - m}{\sigma/\sqrt{n}}\right)$ is Asymptotically normal, if X_i's are independent identical distributions.

Hence Proved.

Q3. Here $\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$ where

$(x_1, x_2, \ldots, x_n)$ is the sample of size n.

$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$, and $\bar{X}$ is the corresponding sampling distribution.

To show $\Rightarrow$ $\bar{X} \sim \text{Normal}\left(m, \frac{\sigma}{\sqrt{n}}\right)$ $\forall n$.

X_i's are independent Normal(m, σ) variates.

By reproductive property

$X_1 + X_2 + \cdots + X_n$ is a Normal$(nm, \sqrt{n\sigma^2})$ variate.

$m_1 + m_2 + \cdots + m_n$, $\sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2}$

Let it be S_n.

$\Rightarrow S_n \sim \text{Normal}(nm, \sqrt{n}\sigma)$

$\Rightarrow \frac{S_n - nm}{\sqrt{n\sigma^2}} \sim \text{Normal}(0, 1)$

$\Rightarrow \frac{\frac{S_n}{n} - m}{\frac{\sqrt{n}\sigma}{\sqrt{n}}} \sim \text{Normal}(0, 1)$

Date
Page

$$\Rightarrow \boxed{\left(\frac{\bar{X}-m}{\sigma/\sqrt{n}}\right) \sim \text{Normal}(0,1)}$$ True $\forall n$. Hence Proved

Q4. Shown that Distribution is $\chi^2(n-1)$.
Now ~~α~~, to show $\Rightarrow \bar{X}, S^2$ are independent.

Consider $Z_1 = a_{11}Y_1 + a_{12}Y_1 + \dots + a_{1n}Y_n$
$Z_2 = a_{21}Y_1 + a_{22}Y_2 + \dots + a_{2n}Y_n$
$\vdots$
$Z_n = a_{n1}Y_1 + a_{n2}Y_2 + \dots + a_{nn}Y_n$.

where each Y_i ~~$\sim$ Norma~~ $= \frac{X_i - m}{\sigma} \sim \text{Normal}(0,1)$.

and each $a_{1i} = \frac{1}{\sqrt{n}}$ $\forall i$. Then $\boxed{Z_i\text{'s are independent}}$

$$Z_1 = \frac{1}{\sqrt{n}}\left(\cancel{(Y_1+Y_2+\dots+Y_n)}\ \frac{(X_1+X_2+\dots+X_n)-nm}{\sigma}\right).$$

$$= \frac{\frac{(X_1+X_2+\dots+X_n)}{n} - \frac{nm}{n}}{\frac{\sigma}{\sqrt{n}}} \cdot \frac{1}{\sqrt{n}} \circ \left(\frac{\bar{X}-m}{\sigma/\sqrt{n}}\right).$$

We can choose the remaining a_{ij} s.t.

$$\sum_{k=1}^{n} a_{ki} a_{kj} = \delta_{ij} = \begin{cases} 1, & i=j \\ 0, & i \neq j \end{cases}$$

Then we have shown that

$$Z_1^2 + \dots + Z_n^2 = Y_1^2 + \dots + Y_n^2.$$

$$\Rightarrow \underbrace{(Y_1^2 + Y_2^2 + \dots + Y_n^2) - Z_1^2}_{\Downarrow\ \left(\frac{nS^2}{\sigma^2}\right)} = Z_2^2 + \dots + Z_n^2.$$

$$\Rightarrow \left(\frac{nS^2}{\sigma^2}\right) = Z_2^2 + \dots + Z_n^2. \Rightarrow \boxed{S^2 = \frac{\sigma^2}{n}\left(\sum_{i=2}^{n} Z_i^2\right)}.$$

~~$\bar{X} = (\bar{X}-m) + (X-m) + (X-m)$~~ & $\boxed{\left(\frac{\sigma}{\sqrt{n}} Z_1 + m\right) = \bar{X}}$

We also know that if $z_1, z_2, \ldots, z_n$ are independent, then any f^n $f(z_1)$ & $g(z_2, \ldots, z_n)$ are independent.

$S^2 = f(z_2, \ldots, z_n)$ and $\bar{x} = g(z_1)$.

$\therefore$ $\boxed{\bar{x} \text{ \& } S^2 \text{ are independent}}$

Q5 Done in Theory.

Q6. To find $\Rightarrow$ Distribution of S^2.

We know that $\frac{nS^2}{\sigma^2} \sim \chi^2(n-1) = X$ (let).

$$\Rightarrow S^2 = \left(\frac{\sigma^2}{n}\right)X.$$

$$f_X(x) = \begin{cases} \frac{1}{2\Gamma(n-1/2)} e^{-\frac{x}{2}} \cdot \left(\frac{x}{2}\right)^{\frac{n-1}{2}-1}, & x > 0. \end{cases}$$

$$S^2 = \frac{\sigma^2}{n} X.$$

$\frac{dS^2}{dX} = \frac{\sigma^2}{n} > 0$. $\therefore$ We can use 1D transformation formula.

$$\therefore f_{S^2}(s^2) = f_X(x) \cdot \left|\frac{dx}{ds^2}\right|$$

$$= \frac{1}{2\Gamma((n-1)/2)} \cdot e^{-\frac{x}{2}} \cdot \left(\frac{x}{2}\right)^{\frac{n-1}{2}-1} \cdot \frac{n}{\sigma^2}$$

$$\Rightarrow \boxed{f_{S^2}(s^2) = \frac{1}{2\Gamma((n-1)/2)} \cdot e^{-\frac{ns^2}{2\sigma^2}} \cdot \left(\frac{ns^2}{2\sigma^2}\right)^{\frac{n-1}{2}-1} \cdot \frac{n}{\sigma^2}, \ s^2 > 0}$$

Q7. (a). Population of X.

where $X \sim$ Binomial (n, p) variate.

Let the sample be of size n.

$(x_1, x_2, \ldots, x_n)$.

Sample mean $\bar{x} = \frac{x_1 + x_2 + \dots + x_k}{k}$.

sampling distribution of $\bar{x} = \bar{X} = \frac{X_1 + X_2 + \dots + X_k}{k}$,

where $X_i \sim$ Binomial (n, p) $\forall i = (1, 2, \dots, k)$.

<u>By Reproductive Property</u>

$X_1 + X_2 + \dots + X_k \sim$ Binomial (nk, p).

$\cancel{\Rightarrow \frac{X_1 + X_2 + \dots + X_k}{k} \sim \text{Binomial}(n, p) \text{ variate}}$

$\Rightarrow k\bar{X} \sim$ Binomial (nk, p)

Let $Y = k\bar{X}$. $\therefore Y \sim$ Binomial (nk, p).

To find $\rightarrow$ distribution of $\bar{X} = \left(\frac{Y}{k}\right)$.

$\cancel{f_{\bar{X}}(\bar{X}) = f_Y(y) \cdot \left|\frac{dy}{d\bar{x}}\right| = f_Y(y) \cdot k}$

$\cancel{\therefore f_{\bar{X}}(\bar{x}) =}$

For discrete distributions.

$f_{\bar{X}}(\bar{x}) = f_Y(y)$ where $\bar{x} = \frac{y}{k}$

$\Rightarrow f_{\bar{X}}(\bar{x}) = f_Y(k\bar{x})$

$= \binom{nk}{y} \cdot (p)^y \cdot (1-p)^{nk-y}$, $y = \{0, 1, \dots, nk\}$

$$\boxed{f_{\bar{X}}(\bar{x}) = \binom{nk}{k\bar{x}} \cdot (p)^{k\bar{x}} \cdot (1-p)^{nk-k\bar{x}}} \Rightarrow \boxed{\bar{x} = \{0, \tfrac{1}{k}, \dots, n\}}$$

PS-1

Q7.(b) $X \sim$ Poisson(μ).

Let $(x_1, x_2, \ldots, x_n)$ is random sample of size n.

$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$, $\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$ is the corresponding sampling distribution.

$X_1, X_2, \ldots, X_n \sim$ Poisson(μ).

By Reproductive Property

$X_1 + X_2 + \ldots + X_n \sim$ Poisson$(n\mu)$.

$\Rightarrow n\bar{X} \sim$ Poisson$(n\mu)$. Let $n\bar{X} = Y$.

$Y \sim$ Poisson$(n\mu)$.

$\Rightarrow \boxed{\bar{X} = \frac{Y}{n}}$ We need to find distribution of $\left(\frac{Y}{n}\right)$.

$f_{\bar{x}}(\bar{x}) = f_Y(y)$ where $y = n\bar{x}$ → for discrete distributions.

$$f_Y(y) = \frac{e^{-n\mu} \cdot (n\mu)^y}{\lfloor y}, \quad y = 0, 1, \ldots \infty.$$

$$\Rightarrow \boxed{f_{\bar{x}}(\bar{x}) = \frac{e^{-n\mu} \cdot (n\mu)^{n\bar{x}}}{\lfloor n\bar{x}}} \Rightarrow \text{Distribution of } \bar{X}.$$

$$\boxed{\bar{x} = 0, \tfrac{1}{n}, \ldots \infty.}$$

(c). $X \sim \text{Gamma}(l)$

$x_1, x_2, \ldots, x_n$ → Random sample of size n.

$$\underbrace{\bar{X} = \left(\frac{x_1 + \ldots + x_n}{n}\right)}_{\text{statistic}} \Rightarrow \underbrace{\boxed{\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}}}_{\text{Sampling Distribution}}$$

By reproductive property.

$X_1 + X_2 + \ldots + X_n \sim \text{Gamma}(nl)$.

To find ⇒ Distribution of $\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = \bar{X}$.

Let $Y = \sum_{i=1}^{n} X_i$. $\therefore Y \sim \text{Gamma}(nl)$.

$$\boxed{\bar{X} = \frac{Y}{n}}.$$

~~$f_{\bar{x}}(\bar{x}) = f_Y(y)$ where $y = n\bar{x}$ for discrete~~

$$f_{\bar{x}}(\bar{x}) = f_Y(y) \cdot \left|\frac{dy}{d\bar{x}}\right| \Rightarrow \text{for continuous case}$$

$$\Rightarrow f_{\bar{x}}(\bar{x}) = f_Y(y) \cdot n$$

$$\Rightarrow f_{\bar{x}}(\bar{x}) = \frac{1}{\Gamma(nl)} \cdot e^{-y} \cdot (y)^{nl-1} \cdot n.$$

$$\Rightarrow f_{\bar{x}}(\bar{x}) = \frac{1}{\Gamma(nl)} e^{-n\bar{x}} \cdot (n\bar{x})^{nl-1} \cdot n$$

$$\Rightarrow \boxed{f_{\bar{x}}(\bar{x}) = \frac{1}{\Gamma(nl)} e^{-n\bar{x}} \cdot (\bar{x})^{nl-1} \cdot (n)^{nl}}$$