UNIT – 5 (Sampling)

Distribution of $\bar{x}$, S^2

Data:- Sampled from a Normal Distribution

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad , \quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

(Sample mean) (unbiased sample variance)

Theorem- 1 let $X_1, X_2, \cdots X_n$ be a random sample

$x_i \sim N(\mu, \sigma^2)$, $i = 1, 2, 3, \cdots n$

then $\bar{x}$ and $x_1 - \bar{x}$, $x_2 - \bar{x}$, $\cdots$ $x_n - \bar{x}$ are independent.

Proof:- we evaluate the MGF of $\bar{x}, x_1 - \bar{x}, x_2 - \bar{x}, \cdots, x_n - \bar{x}$

$$M_x(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

$$M(t, t_1, t_2, \cdots t_n) = E\left[e^{t\bar{x} + t_1(x_1 - \bar{x}) + \cdots + t_n(x_n - \bar{x})}\right]$$

$$= E\left[e^{\sum_{i=1}^{n} t_i x_i - \left(\sum_{j=1}^{n} t_j - t\right)\bar{x}}\right]$$

$$= E\left[e^{\sum_{i=1}^{n} t_i x_i - \left(\sum_{j=1}^{n} t_j - t\right)\frac{1}{n}\sum_{i=1}^{n} x_i}\right]$$

$$= E\left[e^{\sum_{i=1}^{n} x_i \left(t_i - \frac{\left(\sum_{j=1}^{n} t_j - t\right)}{n}\right)}\right] \qquad \bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i$$

$$= E\left[e^{\sum_{i=1}^{n} u_i x_i}\right] \qquad u_i = \frac{nt_i - n\bar{t} + t}{n}$$

$$= \prod_{i=1}^{n} E\left[e^{u_i x_i}\right] \qquad \text{As } x_i \text{ are independent variables.}$$

$$E[e^{tX}] = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

$$\cancel{M(t_1, t_2} \quad M(t, t_1, t_2, \dots t_n) = \prod_{i=1}^{n} e^{\mu u_i + \frac{\sigma^2 u_i^2}{2}}$$

$$= e^{\mu \sum_{i=1}^{n} u_i + \frac{\sigma^2}{2} \sum_{i=1}^{n} u_i^2} \qquad \bar{t} = \frac{1}{n} \sum_{i=1}^{n} t_i$$

$$= e^{\mu \sum_{i=1}^{n} \frac{nt_i - n\bar{t} + t}{n} + \frac{\sigma^2}{2} \sum_{i=1}^{n} \left(\frac{nt_i - n\bar{t} + t}{n}\right)^2}$$

$$= e^{\mu t} \cdot e^{\mu \sum_{i=1}^{n} \cancel{(t_i - \bar{t})} \to 0} \; e^{\frac{\sigma^2}{2n^2} \sum_{i=1}^{n} n^2 (t_i - \bar{t})^2}$$

$$e^{\frac{\sigma^2}{2n^2} n\bar{t}^2} \; e^{\cancel{\frac{\sigma^2}{2n^2} \sum_{i=1}^{n} 2\bar{t}n(t_i - \bar{t})} \to 0}$$

$$M(t, t_1, t_2, \dots t_n) = \underbrace{e^{\mu t + \frac{\sigma^2}{2n} t^2}}_{MGF} \cdot \underbrace{e^{\frac{\sigma^2}{2} \sum_{i=1}^{n} (t_i - \bar{t})^2}}_{M_{X_i}(t_1 - \bar{t}, t_2 - \bar{t}, \dots)}$$

but need $M_{X_1 - \bar{X}, X_2 - \bar{X}, \dots}(t_i)$

$$M_X(t) = E[e^{t\bar{X}}] \qquad \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

$$e^{\frac{\sigma^2}{2} \sum_{i=1}^{n} (t_i - \bar{t})^2} = E\left[e^{\sum_{i=1}^{n} (t_i - \bar{t}) X_i}\right]$$

$$= E\left[e^{\sum_{i=1}^{n} t_i X_i - \bar{t} \sum_{i=1}^{n} X_i}\right]$$

$$= E\left[e^{\sum_{i=1}^{n} t_i X_i - n\bar{X}\bar{t}}\right]$$

$$= E\left[e^{\sum_{i=1}^{n} t_i (X_i - \bar{X})}\right]$$

$$M_{X_1 - \bar{X}, X_2 - \bar{X}, \dots}(t_1, t_2, t_3, \dots t_n)$$

$\equiv$ Separation of MGF's

$\Rightarrow$ Independence

* $\bar{X}$ and s^2 are independent

$\hookrightarrow$ Sum of all $X_i - \bar{X}$

* $\frac{(n-1)s^2}{\sigma^2} \sim \psi^2(n-1)$

chi-square distribution

$$\boxed{\psi^2(K) = f_1(x) = \frac{1}{2^{K/2}\Gamma K/2} x^{k/2-1} e^{-x/2}, \quad x>0}$$

$$\Gamma(\alpha,\beta) \quad f_2(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x>0$$

$$\psi^2(k) = \Gamma\left(\frac{k}{2}, \frac{1}{2}\right)$$

let $X_i \sim N(0,1)$, $Z_i = X_i^2$, $X = \sqrt{Z}$, $\frac{dx}{dt} = \frac{1}{2} z^{-1/2} \frac{dz}{dt}$

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right)$$

$$f(z) = \left|\frac{1}{2} z^{-1/2}\right| f_x(\sqrt{z}) + \left|\frac{1}{2} z^{-1/2}\right| f_x(-\sqrt{z})$$

$$f(z) = \frac{1}{2} z^{-1/2} \frac{1}{\sqrt{2\pi}} e^{-3/2} \cdot 2$$

$$f(z) = \frac{1}{\sqrt{2\pi z}} e^{-3/2}$$

$$f(z) = \left(\frac{1}{2}\right)^{1/2} \frac{1}{\Gamma 1/2} z^{1/2-1} e^{-z/2} = \Gamma\left(\frac{1}{2}, \frac{1}{2}\right)$$

$$\boxed{Y = \sum_{i=1}^{n} Z_i = \sum_{i=1}^{n} X_i^2}$$

$$\psi^2(n) \text{ or } \Gamma\left(\frac{n}{2}, \frac{1}{2}\right)$$

MGF of $\Gamma(\alpha,\beta)$ Gamma distribution

$$\boxed{M(t) = \left(1 - \frac{t}{\beta}\right)^{-\alpha}, \quad t<\beta}$$

$$M_Y(t) = E\left[e^{tY}\right] = E\left[e^{t\sum_{i=1}^{n} z_i}\right]$$

X_i are IID, Z_i are IID

$$M_Y(t) = \prod_{i=1}^{n} E\left(e^{tz_i}\right) = \prod_{i=1}^{n} (1-2t)^{-1/2}$$

Uniqueness of MGF, $M_Y(t) = (1-2t)^{-n/2}, \; t < 1/2$

$$f(y) = \Gamma\left(\frac{n}{2}, \frac{1}{2}\right) = \chi^2(n)$$

Goal: $\dfrac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1)$

Proof:-

$$\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma^2} \sim \chi^2(n) \text{ --- } (1)$$

$$\frac{n(\bar{x}-\mu)^2}{\sigma^2} \sim \chi^2(1) \text{ --- } (2)$$

$\bar{x}$ and s^2 are independent.

$$A = \sum_{i=1}^{n} \frac{(x_i-\mu)^2}{\sigma^2} = \sum_{i=1}^{n} \frac{1}{\sigma^2}(x_i - \bar{x} + \bar{x} - \mu)^2$$

$$A = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2 + \frac{1}{\sigma^2}\sum_{i=1}^{n}(\bar{x}-\mu)^2 + \frac{2}{\sigma^2}\cancelto{0}{\sum_{i=1}^{n}(x_i-\bar{x})}(\bar{x}-\mu)$$

$$A = \frac{1}{\sigma^2}\underbrace{(n-1)s^2}_{B} + \frac{1}{\sigma^2}\underbrace{n(\bar{x}-\mu)^2}_{C}$$

$$M_A(t) = (1-2t)^{-n/2}, \quad t < 1/2$$

$$M_{B+C}(t) = E\left[e^{t\left(\frac{1}{\sigma^2}(n-1)s^2 + \frac{1}{\sigma^2}n(\bar{x}-\mu)^2\right)}\right]$$

As s^2 & $\bar{x}$ are independent

$$M_{B+C}(t) = E\left[e^{t\left(\frac{1}{\sigma^2}(n-1)s^2\right)}\right] \cdot E\left[e^{t\frac{n}{\sigma^2}(\bar{x}-\mu)^2}\right]$$

$$E\left[e^{t\frac{(n-1)}{\sigma^2}s^2}\right] = \frac{(1-2t)^{-n/2}}{(1-2t)^{-1/2}} = (1-2t)^{-\frac{(n-1)}{2}}$$

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

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 , \quad s=\sqrt{s^2}=\frac{1}{\sqrt{n-1}}\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$\Downarrow$

standard deviation $\qquad f_S(s) = ?$

$$\boxed{\begin{gathered}\frac{(n-1)s^2}{\sigma^2} \sim \psi^2(n-1) \\ \frac{\sqrt{n-1}\,s}{\sigma} \sim \psi(n-1)\end{gathered}}$$

$$\alpha = \frac{(n-1)s^2}{\sigma^2} , \quad f(\alpha) = \psi^2(n-1) , \quad \boxed{s = \left(\frac{\sigma^2\alpha}{n-1}\right)^{1/2}}$$

$$f_A(\alpha) = \frac{1}{2^{\frac{n-1}{2}}\Gamma\frac{n-1}{2}}\,\alpha^{\frac{n-1}{2}-1}e^{-\alpha/2} , \quad \alpha>0$$

$$\boxed{f_S(s) = \left|\frac{d\alpha}{ds}\right| f_A\left(\frac{(n-1)s^2}{\sigma^2}\right)}$$

$$\frac{d\alpha}{ds} = \frac{2s(n-1)}{\sigma^2}$$

$$f_S(s) = \frac{2s(n-1)}{\sigma^2}\cdot\frac{1}{2^{\frac{n-1}{2}}\Gamma\frac{n-1}{2}}\left(\frac{n-1}{\sigma^2}s^2\right)^{\frac{n-1}{2}-1}e^{-\frac{(n-1)s^2}{2\sigma^2}} , \quad s>0$$

$$w = \frac{\sqrt{n-1}\cdot s}{\sigma} \Rightarrow s = \frac{\sigma}{\sqrt{n-1}}\cdot w$$

$$f_S(s) = \frac{1}{2^{\frac{n-1}{2}-1}\Gamma\frac{n-1}{2}}\left(\frac{n-1}{\sigma^2}\right)^{\frac{n-1}{2}+1-1}s^{n-1-2+1}e^{-\frac{(n-1)s^2}{2\sigma^2}} , \quad s>0$$

$$\boxed{f_S(s) = \frac{1}{2^{\frac{n-1}{2}-1}\Gamma\frac{n-1}{2}}\left(\frac{n-1}{\sigma^2}\right)^{\frac{n-1}{2}}s^{n-1-1}e^{-\frac{(n-1)s^2}{2\sigma^2}} , \quad s>0}$$

$$f(w) = \psi(n-1)$$

$$f(w) = \frac{1}{2^{\frac{n-1}{2}-1}\Gamma\frac{n-1}{2}}\,w^{n-1-1}e^{-w^2/2} , \quad w>0$$

Std. Dev $\qquad \frac{\sqrt{n-1}}{\sigma}s \sim \psi(n-1)$

$w = \frac{\sqrt{n-1}}{\sigma} s$. $\quad f(w) = \left|\frac{ds}{dw}\right| f_S\left(\frac{\sigma w}{\sqrt{n-1}}\right)$.

$s = \frac{\sigma}{\sqrt{n-1}} w$

$$f(w) = \frac{\sigma}{\sqrt{n-1}} \frac{1}{2^{\frac{n-1}{2}-1}\Gamma\frac{n-1}{2}} \left(\frac{n-1}{\sigma^2}\right)^{\frac{n-1}{2}} \left(\frac{\sigma w}{\sqrt{n-1}}\right)^{n-1-1} e^{-w^2/2}$$

$$\boxed{\begin{aligned} f(w) &= \frac{1}{2^{\frac{n-1}{2}-1}\Gamma\frac{n-1}{2}} w^{n-1-1} e^{-w^2/2}, \quad w>0 \\ &\downarrow \\ \psi(n-1) &= 0, \text{ elsewhere} \end{aligned}}$$

$$Y = \frac{\sqrt{n}}{S}(\bar{x}-\mu) = \frac{\sqrt{n}(\bar{x}-\mu)/\sigma}{\sqrt{(n-1)s^2/\sigma^2/(n-1)}}, \quad Y \sim t(n-1)$$

$$f(y) = \frac{\Gamma\frac{v+1}{2}}{\Gamma\frac{v}{2}\sqrt{v\pi}}\left(1+\frac{y^2}{v}\right)^{-\left(\frac{v+1}{2}\right)}$$

$u = \sqrt{n}(\bar{x}-\mu) \qquad var(\sqrt{n}\,\bar{x}) = n\, var(\bar{x}) = \cancel{n}\frac{\sigma^2}{\cancel{n}} = \sigma^2$

$$f_1(u) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-u^2}{2\sigma^2}\right), \quad u \sim N(0,\sigma^2)$$

$$f_2(v) = g_n v^{n-1-1} e^{-\frac{n-1}{2\sigma^2}v^2}, \quad v>0$$

$$g_n = \frac{1}{2^{\frac{n-1}{2}-1}\Gamma\frac{n-1}{2}}\left(\frac{n-1}{\sigma^2}\right)^{\frac{n-1}{2}} \qquad Y = \frac{\sqrt{n}(\bar{x}-\mu)}{S}$$

$$f_Y(y) = \int |u| f_1(yu) f_2(u)\, du$$

$$f_Y(y) = \int_0^\infty u \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2u^2}{2\sigma^2}} g_n u^{n-1-1} e^{-\frac{n-1}{2\sigma^2}u^2} du$$

$g_n' = \frac{g_n}{\sqrt{2\pi\sigma^2}}$

$$f_Y(y) = g_n' \int_0^\infty u^{n-1} e^{-\frac{u^2}{2\sigma^2}(y^2+n-1)} du$$

$\beta = \frac{y^2+n-1}{\sigma^2}$

$$f_Y(y) = g_n' \int_0^\infty \frac{v^{n-1}}{(\sqrt{\beta})^{n-1}} e^{-\frac{v^2}{2}} \frac{1}{\sqrt{\beta}} dv \qquad \sqrt{\beta}\, u = v$$

$$f(y) = \frac{\partial n'}{(\beta)^{\frac{n-1}{2}+\frac{1}{2}}} \cdot \int_0^\infty v^{n-1} e^{-v^2/2} \, dv$$

$$f(y) = \frac{\left(\frac{n-1}{\sigma^2}\right)^{\frac{n-1}{2}}}{2^{\frac{n-1}{2}-1}\,\Gamma\frac{n-1}{2}} * \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \frac{1}{\left(\frac{y^2+n-1}{\sigma^2}\right)^{n/2}} \, 2^{\frac{n}{2}-1}\,\Gamma\frac{n}{2}$$

$$f(y) = \frac{1}{\sqrt{\pi}} (\sigma^2)^{\cancel{\left(\frac{n-1}{2}\right)-\frac{1}{2}+\frac{n}{2}} \to 0} (n-1)^{\frac{n-1}{2}-\frac{n}{2}} \cdot \frac{\Gamma\frac{n}{2}}{\Gamma\frac{n-1}{2}} \cdot \frac{1}{\left(1+\frac{y^2}{n-1}\right)^{n/2}}$$

$$\boxed{f(y) = \frac{1}{\sqrt{(n-1)\pi}} \, \frac{\Gamma\left(\frac{n-1+1}{2}\right)}{\Gamma\frac{n-1}{2}} \left(1+\frac{y^2}{n-1}\right)^{-\left(\frac{n-1+1}{2}\right)} \longrightarrow t(n-1) \quad \nu = n-1}$$

Student's 't' distribution

Point Estimation

$$P(x, \theta)$$

$$P(x_1, x_2, x_3 \cdots\cdots, \theta_1, \theta_2, \cdots\cdots)$$

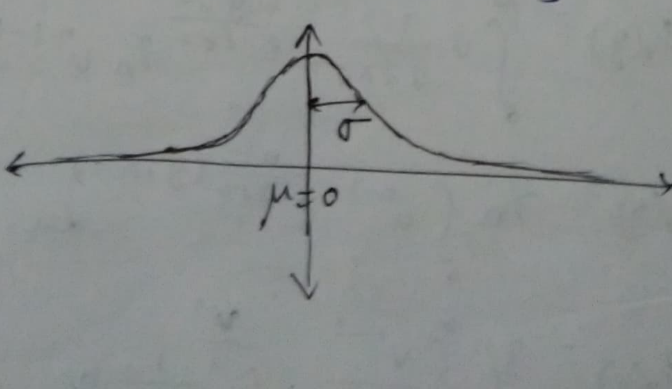
$\downarrow$ vector observation $\downarrow$ parameters.

Def :- Parameter space

The set of all admissable values of the parameters of a distribution function D.F- $F_\theta(x)$ is called parameter space.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) \quad \left\{ \begin{array}{l} -\infty < \mu < \infty \\ 0 < \sigma^2 < \infty \end{array} \right\}$$

$f(x; \mu, \sigma^2)$ $\theta = \{\mu, \sigma^2\}$ parameter space



Estimation problem: Given a random sample $X = \{x_1, x_2, x_3, \cdots x_n\}$ and a paramater vector $\theta = \{\theta_1, \theta_2, \theta_3, \cdots, \theta_k\}$, $k \leq n$

$X \sim F(x)$ Objective: Determine a function $\psi(\theta)$

$\psi(\cdot) \to$ real valued function.

Def: point Estimator: let $X = \{x_1, x_2, \cdots x_n\} \sim P_\theta$ $\quad \theta \to$ parameter vector

A statistic $h(X)$ is said to be a point Estimator of $\psi(\theta)$ if

$h: \mathcal{X} \to \theta$ $\qquad h(X) \to$ approximates $\psi(\theta)$ (estimates)

$\hookrightarrow$ space of data

r.s $X = \{x_1, x_2, \cdots x_n\}$ $\qquad G(\underline{\theta})$

$\theta = \{\theta_1, \theta_2, \cdots \theta_n\}$

Eg. of estimation :-

$$h_1(X) = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$h_2(X) = \frac{2}{n(n+1)} \sum_{i=1}^{n} i\, x_i$$

$$h_3(X) = x_1$$

$$h_4(X) = \frac{x_1 + x_n}{2}$$

$$h_5(X) = \sum_{j=1}^{n} a_j x_j$$

* Infinite number of possible estimators exists.

Q: Optimal Estimation

Choose / Design estimator

optimality criterion, cost fxn, objective fxn

$h(x) \to \psi(\theta)$

$|h(x) - \psi(\theta)| \to$ absolute error.

(i) $P_\theta\{|h(X) - \psi(\theta)| < \varepsilon\} \to$ using probability.

(ii) $E_\theta\{|h(X) - \psi(\theta)|^r\} \to$ using estimation.

max probability approach:- $P_\theta\{|h_0(x) - \psi(\theta)| < \varepsilon\} \geq P_\theta\{|h(x) - \psi(\theta)| < \varepsilon\}$

$h(x)$: h_0, h_1, h_2, $\cdots$, h_n

$h_0 \to$ optimal max $P_\theta(\cdot)$

$$h_0(X) = \arg\max_{\text{w.r.t } h} P_\theta\{|h(X) - \psi(\theta)| < \varepsilon\}$$

<u>Minimum error approach</u>

$$MSE_\theta(h_0) \leq MSE_\theta(h(X))$$

$$\boxed{MSE_\theta(h) = E_\theta\left[(h(X) - \psi(\theta))^2\right]} \rightarrow \text{Mean square error}$$

$$h_0(X) = \arg\min_{\text{w.r.t } h} MSE_\theta(h(X))$$

Variance: $E\left(h(X) - E(h(X))\right)^2$ w.r.t mean of the data

Bias: $b(h, \theta) = h(X) - \psi(\theta)$ w.r.t origin (target)

Avg of all the lengths from target

$$E_\theta\left[h(X) - \psi(\theta)\right]^2 = E_\theta\left\{\left(h(X) - Eh(X) + Eh(X) - \psi(\theta)\right)^2\right\}$$

$$E_\theta\left[h(X) - \psi(\theta)\right]^2 = E_\theta\left(h(X) - Eh(X)\right)^2 + \left(E(h(X)) - \psi(\theta)\right)^2 + 2.E\left[\underbrace{\left(h(X) - E(h(X))\right)}_{0}\left(E(h(X)) - \psi(\theta)\right)\right]$$

$$\boxed{MSE_\theta(h) = var(h) + \left(bias(h, \theta)\right)^2}$$

To get the estimated value closer $\downarrow$ var(h) & $\downarrow$ bias.

<u>Strategy to minimise $MSE_\theta(h)$</u>

Both var(h) and bias need to be reduced.

1) <u>Mininum variance unbiased Estimation</u>

consider unbiased estimation,

$$E_\theta(h(X)) = \psi(\theta) \quad \forall\, \theta \in \Theta$$

On average h(X) estimates $\psi(\theta)$

To minimise MSE_θ (var(h) should also be minimised)

$$\arg\min_h var(h(X))$$

$$h(X) \rightarrow \underbrace{MVUE}$$

minimum variance unbiased estimators.

Examples: consider a Random variable $X \sim N(\mu, \sigma^2)$

Ramdom sample $\{X_1, X_2, \cdots X_n\}$

Show that the following estimator is unbiased.

(i) Sample mean

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad \Rightarrow \quad E(\bar{X}) = \mu \ , \ b(\bar{X}, \mu) = 0$$

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

$$\psi(\theta) = \mu$$

(ii) Sample variance

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

Is $E[S^2] = \sigma^2$? (unbiased or not)

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

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

$$E[S^2] = \frac{1}{n-1}\sum_{i=1}^{n} E\left[X_i^2 + \bar{X}^2 - 2X_i\bar{X}\right]$$

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

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

$$\mathrm{var}(X_i) = E(X_i^2) - (E(X_i))^2$$

$$\sigma^2 = E(X_i^2) - \mu^2$$

$$\boxed{E[X_i]^2 = \sigma^2 + \mu^2}$$

$$E[\bar{X}]^2 = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i \frac{1}{n}\sum_{j=1}^{n} X_j\right]$$

$$E[\bar{X}]^2 = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} E(X_i, X_j)$$

$$E[\bar{X}]^2 = \frac{1}{n^2}\left[\sum_{i=1}^{n} E(X_i^2) + \sum_{i}\sum_{j} E(X_i)E(X_j)\right]$$

$i = j \qquad i \neq j$

$$E[\bar{X}]^2 = \frac{1}{n^2}\left[n(\sigma^2 + \mu^2) + (n^2 - n)\mu^2\right]$$

$$\boxed{E[\bar{X}]^2 = \frac{\sigma^2}{n^2} + \mu^2}$$

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

$$E[s^2] = \frac{1}{n-1}\left[n\sigma^2 + \cancel{n\mu^2} - \sigma^2 - \cancel{n\mu^2}\right]$$

$$E[s^2] = \frac{1}{\cancel{n-1}}\cancel{(n-1)}\sigma^2 \Rightarrow \boxed{E[s^2] = \sigma^2}$$

$\boxed{\text{unbiased variance.}}$