Estimators and "n"

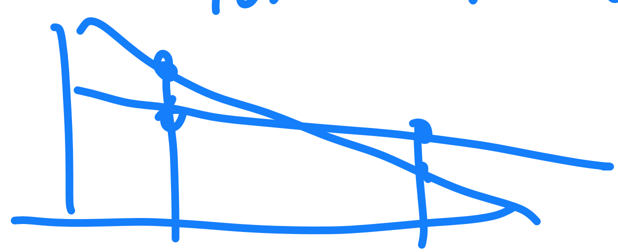
$$\hat{\theta}_n \leftarrow T(X_1, \ldots, X_n)$$

$\hat{\theta}^1_n$, $\hat{\theta}^2_n$?

unbiased

lowest variance - given n

S_n^2



$X_i \sim \mu, \sigma^2$ Sample Variance is Unbiased

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

$$Var(Y) = E[Y^2] - (E[Y])^2$$

$$E[Y^2] = Var[Y] + (E[Y])^2$$

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

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

$$= \frac{1}{n-1} \sum_i \left[\underbrace{Var(X_i - \bar{X}_n)} + \cancel{(E[X_i - \bar{X}_n])^2}\right]$$

0

$$\therefore \quad Var[\underbrace{X_i - \bar{X}_n}]$$

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j$$

$$X_i - \frac{X_i}{n} - \frac{1}{n}\sum_{j\neq i} X_j$$

$$\sum_{j\neq i} X_j = \sum_{j=1}^{j-1} X_j + \sum_{j=i+1}^{n} X_j$$

$$\left(\underbrace{\frac{n-1}{n} X_i} - \underbrace{\frac{1}{n}\sum_{j\neq i} X_j}\right)$$

$$Var\left(\frac{n-1}{n} X_i\right) - Var\left(\frac{1}{n}\sum_{j\neq i} X_j\right) = \frac{(n-1)^2}{n^2} Var(X_i) - \frac{1}{n^2}\sum_{i\neq j} Var(X_j)$$

$$\frac{(n-1)^2}{n^2} \mathrm{Var}[X_i] - \frac{1}{n^2} \sum_{j \neq i} \mathrm{Var}[X_j]$$

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

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

$$\Rightarrow E[S_n^2] = \sigma^2$$

Confidence Intervals

Definition

R.S. $(X_1, \ldots, X_n)$ $\quad$ F dist

parameter θ

$100(1-\overset{.05}{\alpha})\%$ confidence interval L_n, U_n

$$P(L_n < \theta < U_n) = 1-\alpha$$

Confidence Intervals for Mean

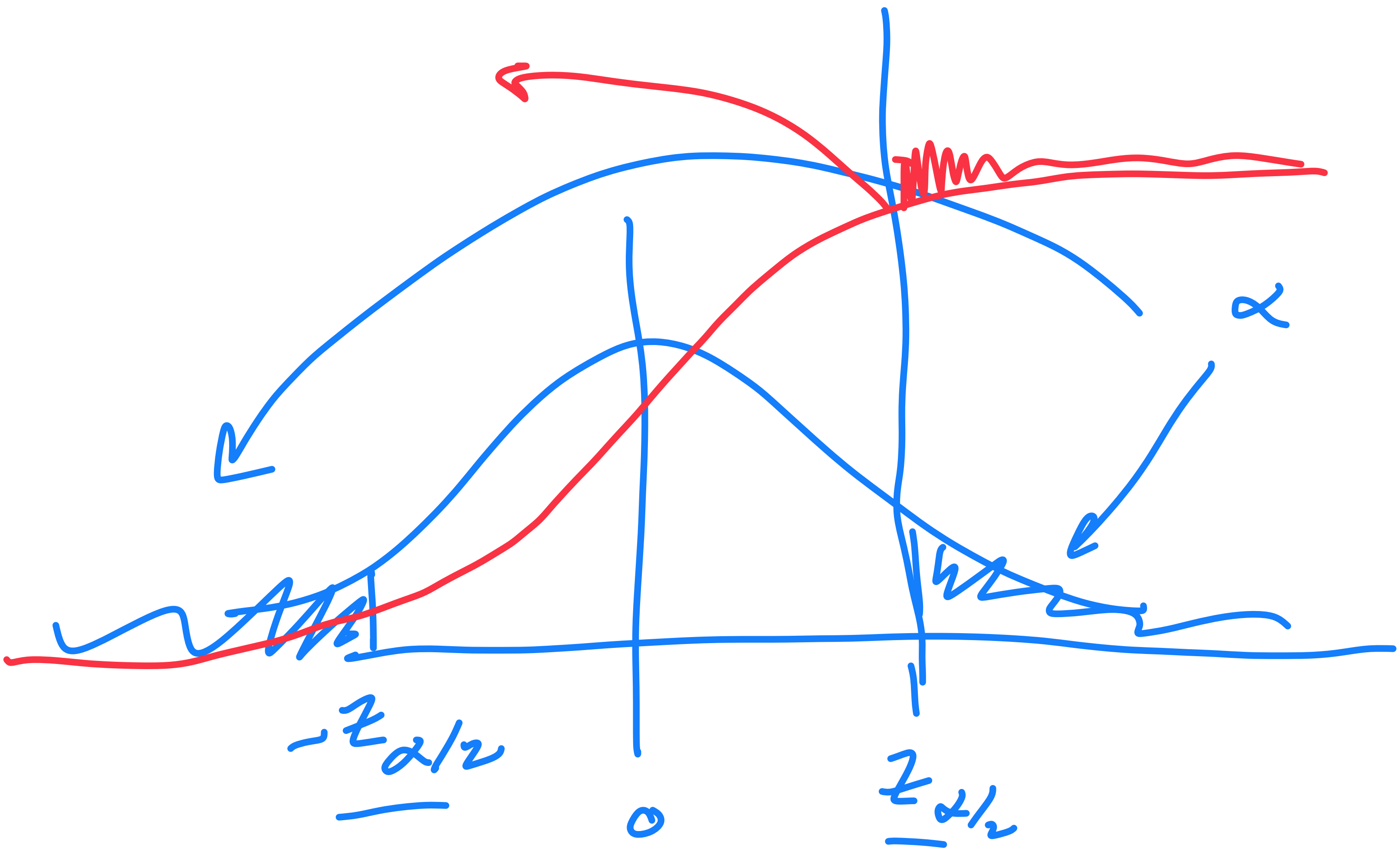
(Known Variance).

$X_i \sim N(\mu, \sigma^2)$ $\qquad X_1 \, .. \, X_n$ $\qquad$ don't know the mean

$\bar{X}_n \rightarrow$ $$Z_n = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

$Z_n \sim N(0,1)$

$$P\left(-z_{\alpha/2} < Z_n < z_{\alpha/2}\right) = 1 - \alpha$$

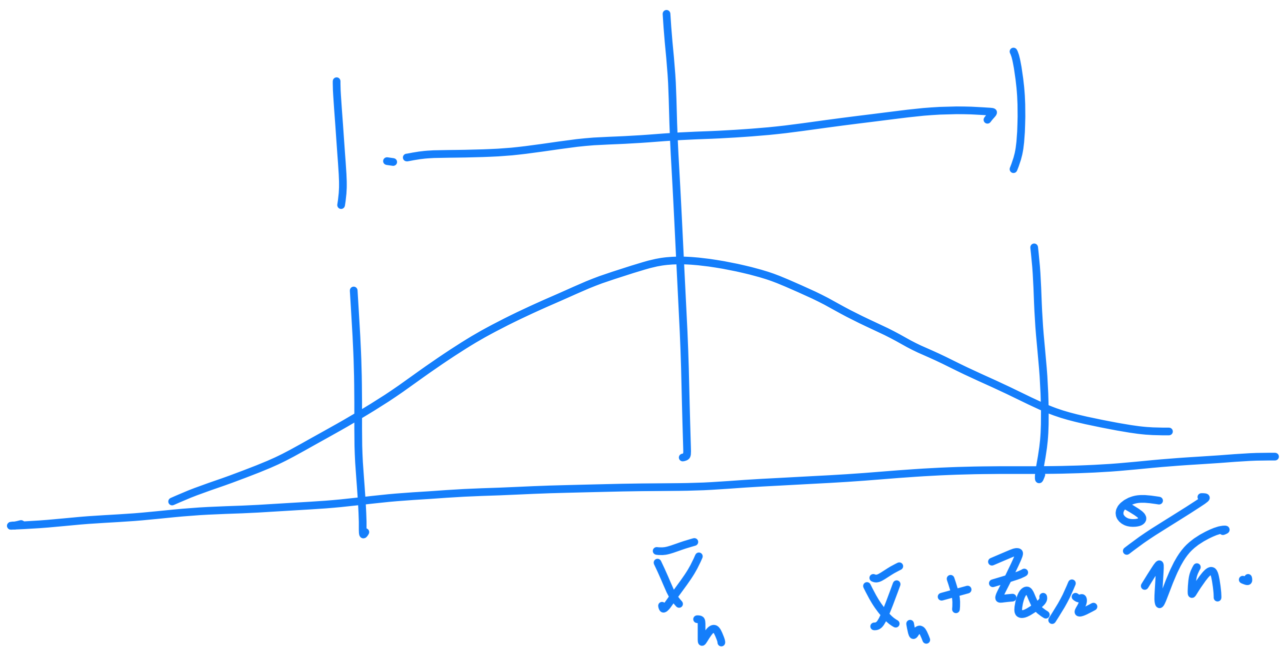


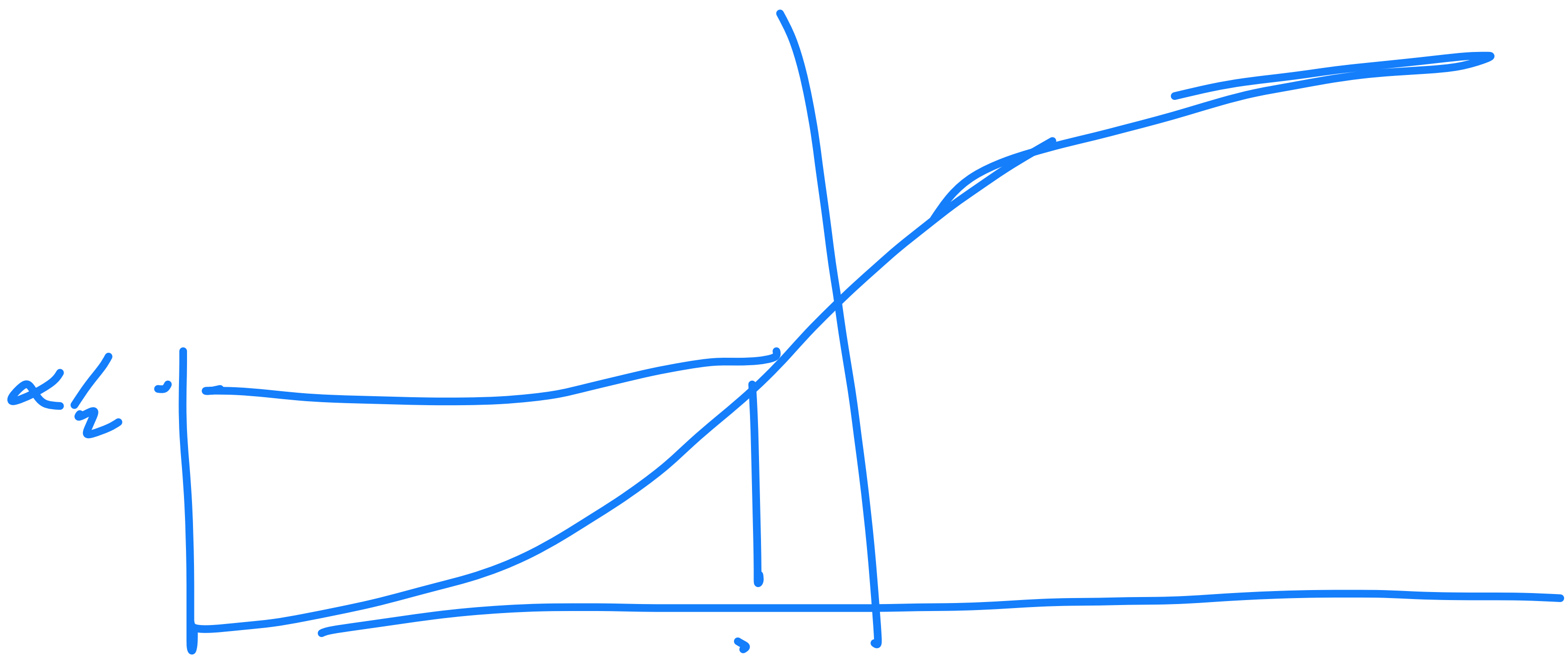
α
$-z_{\alpha/2}$
0
$z_{\alpha/2}$

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

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

$z_{\alpha/2}$ ← qnorm$(1 - \frac{1}{2}\alpha)$




$\alpha/2$

Snowfall in Wasatch

- 40 samples in mountains $\alpha/2$

Sample mean $\rightarrow$ 620 $\alpha = 1 - .95$

Known from previous years var = 36 $= \sigma^2$

Q: what is the 95% confidence in this estimate

$$L_n = 620 - \underset{1.96}{Z_{.025}} \frac{6}{6.3} \leftarrow \frac{\sigma}{\sqrt{n}} = 618...$$

$$U_n = 620 + 1.87 = 621.87.$$

$$[618.13,\ 621.87]$$