Review: Point estimators:

X_i iid f_θ. θ unknown parameter(s).

MLE: maximum likelihood estimator:

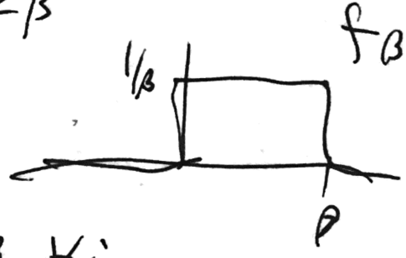
$$L(x_1,\dots,x_n;\theta) = f_\theta(x_1)\cdots f_\theta(x_n).$$

Find θ s.t. $L(\theta)$ is maximal, $\to \hat{\theta}$.

$\Leftrightarrow \log L(\theta)$ maximal.

Sometimes use calculus, $\frac{\partial}{\partial\theta}\log L(\theta) = 0$ solve,
NOT panacea, have to think, if L has boundary conditions.

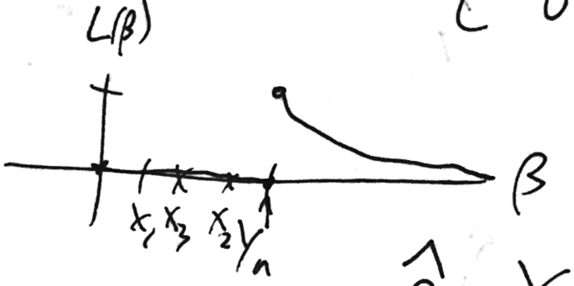
E.g.: $$f_\beta(x) = \begin{cases} 1/\beta, & 0<x<\beta \\ 0 & \text{else} \end{cases}$$



$$L(X_1,\dots,x_n;\beta) = \begin{cases} 1/\beta^n, & 0<X_i<\beta \ \forall i. \\ 0, \end{cases}$$

$\Leftrightarrow \beta > \max X_n = Y_n$

Y_n: nth order statistic.



$$\hat{\beta} = Y_n.$$

Interval estimation:

① $X_1, \dots, X_n$ iid $\sim N(\mu, \sigma^2)$. Want μ.

If σ^2 known: (Point estimators: $\frac{1}{2}(X_1+X_2) = \bar{X} = \hat{\mu}_1$, $2X_1 - X_2 = \hat{\mu}_2$).

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

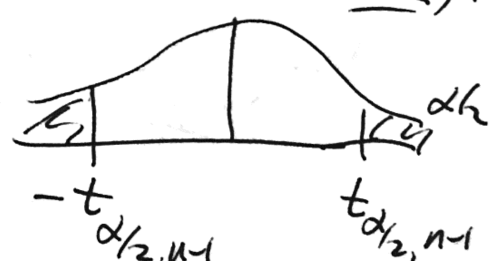
$$\mathbb{P}\left(\left| \frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}} \right| \le z_{\alpha/2} \right) = 1-\alpha$$

$$\Updownarrow$$

$$\bar{X} - \qquad < \mu < \bar{X} + z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}.$$

If σ^2 unknown, $\frac{\bar{X}-\mu}{\sqrt{S^2/n}} \sim T$ distr with $n-1$ deg.

$$\mathbb{P}\left(\left| \frac{\bar{X}-\mu}{\sqrt{S^2/n}} \right| < t_{\alpha/2, n-1} \right) = 1-\alpha.$$



$$\bar{X} - \qquad < \mu < \bar{X} + t_{\alpha/2, n-1}\sqrt{\frac{S^2}{n}}.$$

What if X_i iid normal, Know μ, want σ.

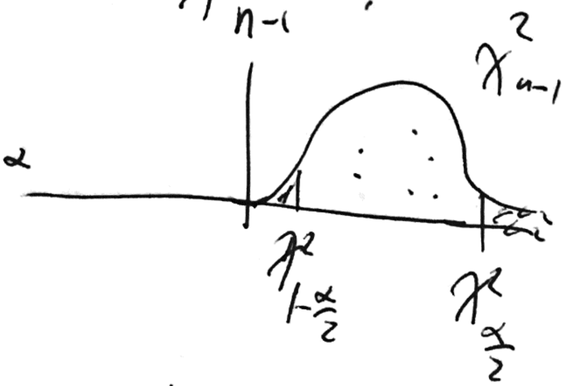
$$\mathbb{P}\left(\underbrace{\left|\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}\right| < z_{\alpha/2}}\right) = 1-\alpha.$$

$$\frac{\sigma^2}{\cancel{n}} \quad \cancel{\sqrt{\frac{\sigma}{n}}} > \left(z_{\alpha/2}\frac{1}{|\bar{X}-\mu|}\right)^2 \cdot n \qquad (\text{silly})$$

Instead:

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

$$\mathbb{P}\left(\underbrace{\chi^2_{1-\frac{\alpha}{2},n-1} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2,n-1}}\right) = 1-\alpha$$



$$\Downarrow$$

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}} < \sigma^2 < S^2(n-1)\frac{1}{\chi^2_{1-\alpha/2,n-1}}$$

---

② $X_1^{(A)}, \ldots, X_{n_A}^{(A)}$, $X_1^{(B)}, \ldots, X_{n_B}^{(B)}$ iid

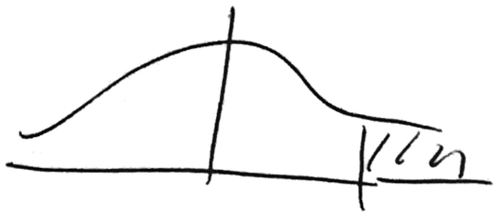
$\sim N(\mu_A, \sigma_A^2)$ $\quad\sim N(\mu_B, \sigma_B^2)$.

Know: σ_A^2, σ_B^2, want $\mu_A-\mu_B$.

$\bar{X}_A - \mu_A \sim N\left(0, \frac{\sigma_A^2}{n_A}\right) \qquad -(\bar{X}_B - \mu_B) \sim N\left(0, \frac{\sigma_B^2}{n_B}\right)$

$$\frac{(\bar{X}_A - \mu_A) - (\bar{X}_B - \mu_B)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \sim N\Big(0, \underset{1}{\cancel{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}}\Big).$$

$$P\Big(\underbrace{\sim\sim\sim < z_\alpha}\Big) = 1-\alpha$$



$$\bar{X}_A - \bar{X}_B - z_\alpha \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} < \mu_A - \mu_B \quad \cancel{\bar{X}_A \pm \bar{X}_B} \leftarrow \text{lower bound.}$$

For upper bound, $P(\sim\sim > -z_\alpha)$.

$$\Rightarrow \mu_A - \mu_B < \bar{X}_A - \bar{X}_B + z_\alpha \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}.$$

If σ_A^2, σ_B^2 unknown, what to do?

$$\frac{\bar{X}_A - \bar{X}_B - (\mu_A - \mu_B)}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \sim Z \sim N(0,1).$$

$$\frac{(n_A - 1) S_A^2}{\sigma_A^2} + \frac{(n_B - 1) S_B^2}{\sigma_B^2} \sim \chi^2_{n_A + n_B - 2}.$$

Stuck unless $\sigma_A = \sigma_B = \sigma$ ← Assume.

$T = \frac{Z}{\sqrt{Y/n}}$

$= Z \cdot \sqrt{\frac{n}{Y}}$.

$$\underbrace{\frac{\bar{X}_A - \bar{X}_B - (\mu_A - \mu_B)}{\cancel{\sigma}\sqrt{\frac{1}{n_A} + \frac{1}{n_B}}} \cdot \sqrt{\frac{(n_A + n_B - 2)}{\frac{1}{\cancel{\sigma^2}}\left((n_A - 1) S_A^2 + (n_B - 1) S_B^2\right)}}}$$

$$\sim T_{n_A + n_B - 2}.$$

Want $(1-\alpha)100\%$ confidence interval upper bound for $\mu_A - \mu_B$. for the

$$P\left(\underbrace{\sim\sim > -t_{\alpha, n_A + n_B - 2}}\right) = 1 - \alpha$$

$$\hookrightarrow \mu_A - \mu_B < \bar{X}_A - \bar{X}_B + t_{\alpha, n_A + n_B - 2} \sqrt{\frac{(n_A - 1) S_A^2 + (n_B - 1) S_B^2}{n_A + n_B - 2}} \sqrt{\frac{1}{n_A} + \frac{1}{n_B}}$$

One sided χ^2. Lower bound

$$X_1, \dots, X_n \sim N(\mu, \sigma^2).$$

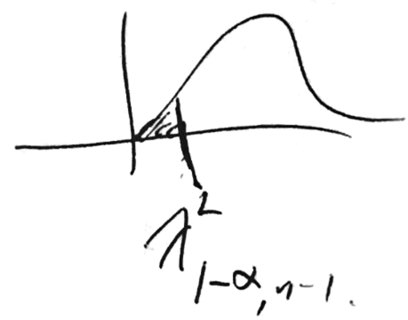
Estimate σ^2 from ~~above~~ below by interval.

$$P\Big(\underbrace{(n-1)\frac{S^2}{\sigma^2} < \chi^2_{\alpha, n-1}}\Big) = 1-\alpha$$

$$\frac{(n-1)\, S^2}{\chi^2_{\alpha, n-1}} < \sigma^2$$

Estimate σ^2 upper bound.

$$1-\alpha = P\Big(\underbrace{\frac{(n-1)\, S^2}{\sigma^2} > \chi^2_{1-\alpha, n-1}}\Big)$$



$$\frac{(n-1)\, S^2}{\chi^2_{1-\alpha, n-1}} > \sigma^2$$

---

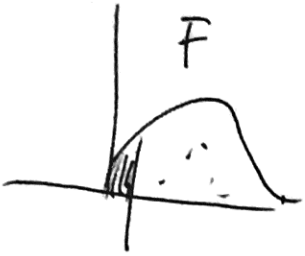
3 $X_1^{(A)}, \dots, X_{n_A}^{(A)}$, $X_1^{(B)}, \dots, X_{n_B}^{(B)}$ want: $\frac{\sigma_A^2}{\sigma_B^2}$.

upper bound.

$$(n_A - 1)\frac{S_A^2}{\sigma_A^2} \sim \chi^2_{n_A-1}, \quad \cdots \text{(B)s.}$$

$$F = \frac{U/n}{V/m}, \quad U, V \text{ indep } \chi^2 \text{ w/ } n, m \text{ (resp.)}$$

degrees of freedom.



$f_{1-\alpha,n,m}$

$$\mathbb{P}\left(\underline{\frac{\cancel{(n_A-1)}\frac{S_A^2}{\sigma_A^2} \Big/ \cancel{(n_A-1)}}{\cancel{(n_B-1)}\frac{S_B^2}{\sigma_B^2} \Big/ \cancel{(n_B-1)}} > f_{1-\alpha,n_A-1,n_B-1}}\right) = 1-\alpha.$$

$$\frac{1}{f_{1-\alpha,n_A-1,n_B-1}} \cdot \frac{S_A^2}{S_B^2} > \frac{\sigma_A^2}{\sigma_B^2}$$

---

(4) $X_1, \ldots, X_n$ iid Bernoulli, $\mathbb{P}(X=1) = \theta$.

$n\bar{X}$ = binomial, $\mathbb{E}(n\bar{X}) = n\theta$, $\mathrm{Var}(n\bar{X}) = n\theta(1-\theta)$.

CLT: $$\underbrace{\mathbb{P}\left(\left|\frac{n\bar{X} - n\theta}{\sqrt{n\theta(1-\theta)}}\right| \leq z_{\alpha/2}\right)} \approx 1-\alpha.$$

can do algebra to get interval in θ.

<u>OR</u>: $$\bar{X} - \quad < \theta < \bar{X} + \underbrace{z_{\alpha/2}\sqrt{\frac{\theta(1-\theta)}{n}}}$$

$$z_{\alpha/2}\sqrt{\frac{\bar{X}(1-\bar{X})}{n}}.$$

(5) $X_1^{(A)}, \dots, X_{n_A}^{(A)}, \quad X_1^{(B)}, \dots, X_{n_B}^{(B)}$

$$\bar{X}_A - \theta_A \approx N\left(0, \frac{\theta_A(1-\theta_A)}{n_A}\right) \approx N\left(0, \frac{\bar{x}_A(1-\bar{x}_A)}{n_A}\right).$$

Same with B's.

$$\mathbb{P}\left(\left|\frac{(\bar{X}_A-\theta_A)-(\bar{X}_B-\theta_B)}{\sqrt{\frac{\bar{x}_A(1-\bar{x}_A)}{n_A}+\frac{\bar{x}_B(1-\bar{x}_B)}{n_B}}}\right| < z_{\alpha/2}\right) = 1-\alpha.$$

$$C: \bar{X}_A-\bar{X}_B - \ < \theta_A-\theta_B < \bar{X}_A-\bar{X}_B + z_{\alpha/2}\left[\sqrt{\cdots}\right].$$

"proportion" = $\frac{\text{Binomial}}{\text{\# trials}}$

Hypothesis testing: H_0, H_1, C critical region, simple / composite, Type I error = C & H_0.

Type II error = C^c & H_1

	H_0 true	H_0 false
C	Type I	
C^c		Type II

Power: $\mathbb{P}(C; H_1) = 1-\beta$

size: $\mathbb{P}(C; H_0) = \alpha$.

<u>Neyman-Pearson Lemma</u>: If C of size α s.t. $\forall \vec{x} \in C$, $L_0(\vec{x}) \le k L_1(\vec{x})$, & $\forall \vec{x} \notin C$, $L_0(\vec{x}) \ge k L_1(\vec{x})$.

$\Rightarrow C$ is a most powerful region for testing H_0 vs H_1.