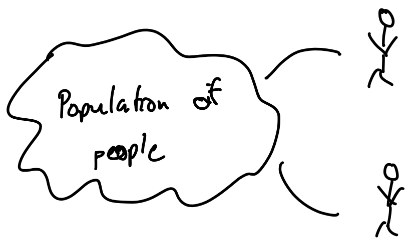
Treatment (Action 1), Control (Action 2)

After taking an action patients die, or live

$X_{i,s} = \mathbb{1}\{\text{the } s\text{th person I gave action } i \text{ lived}\}$



may always respond to drug

may never respond to the drug

Assume people arrive IID $\Rightarrow \mathbb{E}[X_{i,s}] = \theta_i \in [0,1] \quad \forall s$

We say Action 1 is better than action 2

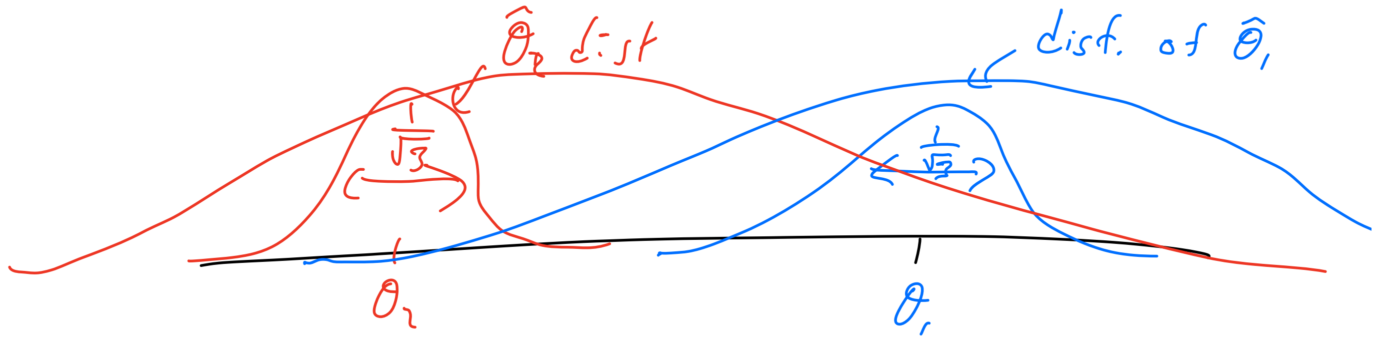
if $\theta_1 > \theta_2$.

Measure action 1, 2 each T times

and compute empirical mean $\hat{\theta}_i = \frac{1}{T} \sum_{t=1}^{T} X_{i,t}$

Question: How big does T need to be

s.t. $\hat{\theta}_1 > \hat{\theta}_2 \implies \theta_1 > \theta_2$



$$\text{Variance}(\hat{\theta}_i) = \mathbb{E}[\hat{\theta}_i^2] - \mathbb{E}[\hat{\theta}_i]^2$$

$$= \frac{\theta_i(1-\theta_i)}{\xi} \leftarrow \in [0, 1/4]$$

Choose ξ to be large enough s.t.

$$\frac{1}{\sqrt{\xi}} \eqsim \max_i \sqrt{\text{Variance}(\hat{\theta}_i)} \leq |\theta_1 - \theta_2|$$

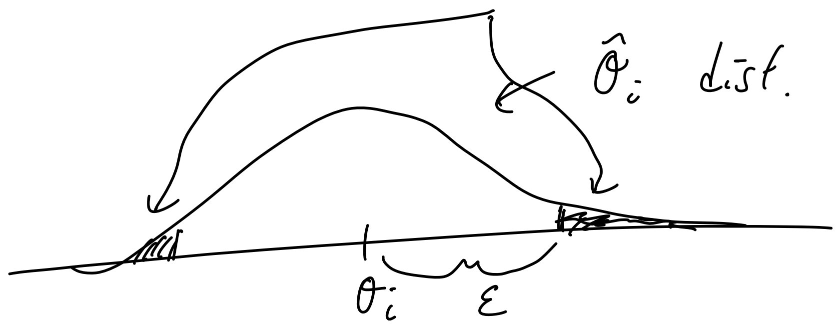
$$\implies \text{pick} \quad \xi \eqsim \frac{1}{|\theta_1 - \theta_2|^2}$$

<u>Defn)</u> $x \lesssim y$ means $\exists$ absolute constant $c > 0$ s.t. $x \leq cy$.

To quantify uncertainty we need a high prob bound:

For $\varepsilon > 0$, $\delta \in (0,1)$ we want to say

$$\mathbb{P}(|\hat{\theta}_i - \theta_i| > \varepsilon) \leq \delta.$$



If $\theta_1 > \hat{\theta}_1 - \varepsilon > \hat{\theta}_2 + \varepsilon > \theta_2$ the with probability at least $1 - \mathbb{P}\left(\bigcup_{i=1}^{2} \{|\hat{\theta}_i - \theta_i| > \varepsilon\}\right)$

$$\geq 1 - \sum_{i=1}^{2} \mathbb{P}(|\hat{\theta}_i - \theta_i| > \varepsilon)$$

$$\geq 1 - 2\delta$$

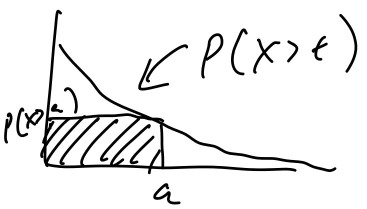
$$\theta_1 > \theta_2$$

Let $Z_1, \ldots, Z_n$ be IID random variables w/ $\mathbb{E}[Z_1] = \mu$, $Z_i \in [0,1]$ $\forall i$

$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} Z_i$ Want ε : $\mathbb{P}(|\hat{\mu}_n - \mu| > \varepsilon) \leq \delta$.

Markov's Inequality For any positive random variable X, $\mathbb{P}(X > t) \leq \frac{\mathbb{E}[X]}{t}$.

Proof $\mathbb{E}[X] = \int_{t=0}^{\infty} \underbrace{\mathbb{P}(X>t)}\,dt$ ← non-increasing in t.



$$\geq a\,\mathbb{P}(X > a) \quad \blacksquare$$

Chebychev's inequality Let X be a R.V w/ $\mathbb{E}[X] = \mu$ and $\mathbb{E}[(X-\mu)^2] = \sigma^2$ then

$$\mathbb{P}(|X-\mu| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

(from above $\mu = \mathbb{E}[\hat{\mu}_n]$, $\sigma^2 \leq \frac{1}{4n}$)

$$\mathbb{P}(|\hat{\mu}_n - \mu| > \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} = \frac{1}{4n\varepsilon^2} = \delta$$

Solve for ε: $\varepsilon = \sqrt{\frac{1}{4n\delta}}$

$$\mathbb{P}\left(|\hat{\mu}_n - \mu| > \sqrt{\frac{1}{4n\delta}}\right) \leq \delta \qquad \text{for any } \delta \in [0,1].$$

At home: construct R.V. X s.t. Chebychev is tight.

Very loose b/c if $Z \sim \mathcal{N}(\mu, \sigma^2)$

$$\mathbb{P}(|Z - \mu| > \varepsilon) \leq \exp(-\varepsilon^2/2\sigma^2)$$

Defn] We say a mean-0 R.V. Z is σ^2-sub-Gaussian if for all $\lambda \in \mathbb{R}$

$$\mathbb{E}[\exp(\lambda Z)] \leq \exp(\sigma^2 \lambda^2/2).$$

Ex. If $Z \sim \mathcal{N}(0, \sigma^2)$ then Z is σ^2-sub-Gaussian.

Warning The sub-Gaussian parameter need not be bounded by the variance in general!

Ex. If Z has PDF $\frac{1}{2}e^{-|x|}$ (i.e Laplace dist.) the variance is $O(1)$ but there does not exist $\sigma^2 > 0$

s.t. Z is σ^2-sub-Gaussian R.V. (confirm at home).

Hoeffding's Lemma | If $Z \in [a,b]$ and $\mathbb{E}[Z]=0$ the Z is (b-a)^2/4 -sub-Gaussian.

Proof sketch: (hard way) Start w/ defn of sub-Gaussian and apply Jensen's inequality.

(easy way) Clever construction of a related R.V. see notes.

Chernoff Bound Technique | Let $Z_1, \ldots, Z_n$ be iid σ^2-sub-Gaussian R.V. w/ $\mathbb{E}[Z_i]=\mu$ then

$$\mathbb{P}\left(\frac{1}{n}\sum Z_i - \mu > \varepsilon\right) \le \inf_{\lambda} \exp(-\lambda\varepsilon)\,\mathbb{E}[\exp(\lambda Z_1)]^n$$

$$\le \exp(-n\varepsilon^2/2\sigma^2).$$

Proof | For any $\lambda > 0$

$$\mathbb{P}\left(\sum Z_i - \mu > n\varepsilon\right) = \mathbb{P}\left(\exp\left(\lambda \sum_{i=1}^{n}(Z_i-\mu)\right) > \exp(\lambda\varepsilon n)\right)$$

(Markov) $$\le \frac{\mathbb{E}\left[\exp\left(\lambda \sum_{i=1}^{n}(Z_i-\mu)\right)\right]}{\exp(\lambda\varepsilon n)}$$

$$= e^{-\lambda n\varepsilon}\,\mathbb{E}\left[\prod_{i=1}^{n}\exp(\lambda(Z_i-\mu))\right]$$

(indep.) $= e^{-\lambda n \varepsilon} \prod_{i=1}^{n} \mathbb{E}[\exp(\lambda(z_i - \mu))]$

(identically dist) $= e^{-\lambda n \varepsilon} \mathbb{E}[\exp(\lambda(z_1 - \mu))]^n$

(sub-Gauss) $\leq e^{-\lambda n \varepsilon} \exp(\lambda^2 \sigma^2 / 2)^n$

$$= \exp(-\lambda n \varepsilon + n\lambda^2 \sigma^2 / 2)$$

$$\frac{\partial}{\partial \lambda}(-\lambda n \varepsilon + \lambda^2 \sigma^2/2) = -n\varepsilon + n\sigma^2 \lambda = 0 \Rightarrow \lambda = \frac{\varepsilon}{\sigma^2}$$

$$\rightarrow \quad = \exp\left(-\frac{n\varepsilon^2}{\sigma^2} + \frac{n\varepsilon^2}{2\sigma^2}\right)$$

(min λ) $= \exp(-n\varepsilon^2 / 2\sigma^2)$

If X and Y are indep R.V. then $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$