Example. $Y_1, Y_2, \ldots, Y_n$ is a random sample from $E(\tau)$ w/ τ unknown.

$$\bar{Y} = \frac{1}{n}(Y_1 + Y_2 + \cdots + Y_n) \sim ?$$

Each $Y_i \sim \Gamma(k=1, \tau)$ and they're independent.

We recall $Y_1 + \cdots + Y_n \sim \Gamma(\underset{\text{shape}}{n}, \underset{\text{scale}}{\tau})$

$\Rightarrow$ $\boxed{\bar{Y} \text{ is NOT A PIVOTAL QUANTITY!}}$

The second parameter of a Γ dist'n is a scale parameter, as we know from

$$\boxed{m_Y(t) = (1 - \tau t)^{-k}} \quad \text{for } t < \frac{1}{\tau}$$

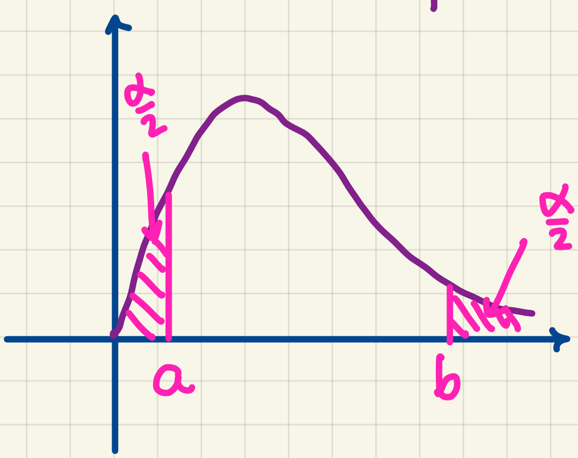
for $Y \sim \Gamma(k, \tau)$

$$m_{aX}(t) = \mathbb{E}\left[e^{aX \cdot t}\right] = \mathbb{E}\left[e^{at \cdot X}\right] = m_X(at)$$

$$\Rightarrow \boxed{U = \frac{1}{\tau}\bar{Y}} = \frac{1}{\tau} \cdot \frac{1}{n} \overbrace{(Y_1 + \cdots + Y_n)}^{\Gamma(n, \tau)}$$

$$\sim \Gamma\left(n, \frac{\cancel{\tau}}{n\cancel{\tau}}\right) = \Gamma\left(n, \frac{1}{n}\right)$$

The dist'n doesn't depend on τ, so we do have a pivotal quatity.



Pick a confidence level. Say $C = 0.90$, i.e., $\alpha = 0.10$.

a = qgamma (0.05, shape = n, scale = 1/n) ←

b = qgamma (0.95, shape = n, scale = 1/n) ←

We know that

$$\mathbb{P}\left[a \leq U \leq b \right] = 0.90$$

$$\mathbb{P}\left[a \leq \frac{1}{\tau} \cdot \bar{Y} \leq b \right] = 0.90$$

$$\mathbb{P}\left[\frac{a}{\bar{Y}} \leq \boxed{\frac{1}{\tau}} \leq \frac{b}{\bar{Y}} \right] = 0.90$$

$$\mathbb{P}\left[\underbrace{\boxed{\frac{\bar{Y}}{b}}}_{\hat{\theta}_L} \leq \tau \leq \underbrace{\boxed{\frac{\bar{Y}}{a}}}_{\hat{\theta}_R} \right] = 0.90$$

$\square$

Approximate Confidence Intervals for p.

Consider a population in which a specific trait occurs w/ an unknown probability p.

Let $(Y_1, Y_2, \dots, Y_n)$ be a random sample from the Bernoulli dist'n w/ the unknown p.

Goal: Designing a confidence interval for p.

Idea: Look @ the natural point estimator for p.

$$\rightarrow \quad \bar{Y} = \boxed{\frac{1}{n}} \Big(\underbrace{Y_1 + Y_2 + \dots + Y_n}_{\sim B(n,p)} \Big)$$

Note: $\mathbb{E}[\bar{Y}] = p$ Of course, the dist'n of $\bar{Y}$ depends on p.

We need a pivotal quantity.

By de Moivre-Laplace Thm:

Set $S_n = Y_1 + \dots + Y_n$

Then, $$\boxed{\frac{S_n - np}{\sqrt{np(1-p)}}} \xRightarrow{\mathcal{D}} N(0,1)$$

We note $\bar{Y} = \frac{1}{n} S_n$

$$U = \frac{\bar{Y} - p}{\sqrt{\frac{p(1-p)}{n}}} \xRightarrow{\mathcal{D}} N(0,1)$$ is still not a pivotal quantity

But we can still create an approximate confidence interval based on it!

Say, C is a confidence level.

Let $z^* = \Phi^{-1}\left(\frac{1+C}{2}\right)$ = qnorm((1+C)/2)

$$\mathbb{P}[-z^* \le U \le z^*] \approx C$$

dM·L

$$\mathbb{P}\left[-z^* \le \frac{\bar{Y}-p}{\sqrt{\frac{p(1-p)}{n}}} \le z^*\right] \approx C$$

$$\mathbb{P}\left[\bar{Y} - z^*\sqrt{\frac{p(1-p)}{n}} \le p \le \bar{Y} + z^* \cdot \sqrt{\frac{p(1-p)}{n}}\right] \approx C$$

Usually, we write $\boxed{\bar{Y} = \hat{p}}$

We construct the confidence interval for p as

$$\boxed{p = \hat{p} \pm z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}$$

Problem 16.2. Gallup's inaugural measure of global loneliness shows over one in five people worldwide (23%) said they felt loneliness "a lot of the day yesterday."[1] *However, there were considerable variations between countries. For instance, out of 1000 individuals polled in Taiwan, 11% reported having felt loneliness "a lot of the day" before. What 90%-confidence would you report for the population proportion of Taiwanese who had felt lonely the day before?*

$\longrightarrow$: $\hat{p} = 0.11$

$$z^* = \Phi^{-1}(0.95) = 1.645$$

$$p = \hat{p} \pm \boxed{z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} = 0.11 \pm 1.645 \cdot \sqrt{\frac{0.11 \cdot 0.89}{1000}}$$

$\square$

Example. How do we figure out the necessary sample size w/ a required margin of error m?

$\longrightarrow$: $n = ?$

$$\boxed{z^* \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \le m} \quad /^2$$

w/ z^* the critical value for the confidence level C

$$(z^*)^2 \cdot \frac{\hat{p}(1-\hat{p})}{n} \le m^2$$

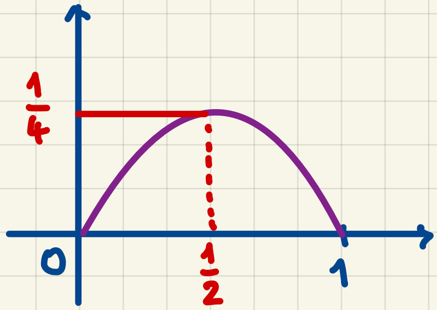
$$(z^*)^2 \cdot \frac{\boxed{\hat{p}(1-\hat{p})}}{m^2} \le n$$

We cannot know $\hat{p}$ before knowing n.

---

[1] https://news.gallup.com/poll/646718/people-worldwide-feel-lonely-lot.aspx

Consider $\hat{p}(1-\hat{p})$ as a function of $\hat{p}$.



The conservative choice for $\hat{p}$ is $\frac{1}{2}$.

The conservative choice for the sample size

$$\boxed{n \geq \left(\frac{z^*}{2m}\right)^2}$$

$\square$