12/1/16

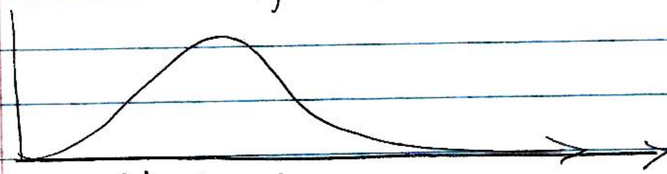
CLT If $X_1 \ldots X_n \overset{iid}{\sim}$ w/ mean μ and s.e σ

If n large

II $\dfrac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \overset{d}{\approx} Z \sim N(0,1)$

III $\bar{X} \approx N(\mu, (\frac{\sigma}{\sqrt{n}})^2)$

IV $T \sim N(n\mu, (\sigma\sqrt{n})^2)$



$\mu = 1000hr$

$\sigma = 500hr$

You buy 50 lightbulbs

$P(\bar{X} \geq 1300hr) =$

$P\left(\dfrac{\bar{X}-1000}{\frac{500}{\sqrt{50}}} > \dfrac{1300-1000}{\frac{500}{\sqrt{50}}}\right)$

$\overset{II}{\approx} P(Z > 4.24) \approx 0$

example Shipments are late 2% of the time. In 10,000 orders, what is than 3% are late?

$X_1 \ldots X_{10,000} \overset{iid}{\sim} Bern(2\%)$

$P(\hat{p} > 3\%) = P\left(\dfrac{\hat{p}-.02}{.0014} > \dfrac{.03-.02}{0.0014}\right)$

$\overset{II}{\approx} P(Z > 7.14) \approx 0$

simple proposition

IV $\bar{X} \sim N(\mu, (\frac{\sigma}{\sqrt{n}})^2)$

$\hat{p} \sim N\left(p, \sqrt{\dfrac{p(1-p)}{n}}\right)$

$N\left(.02, \sqrt{\dfrac{.02(.98)}{10,000}}\right)$

$\mu = p$

$\sigma = \sqrt{p(1-p)}$

$\hat{p} = \bar{X} = \frac{1}{n}\Sigma X_i$

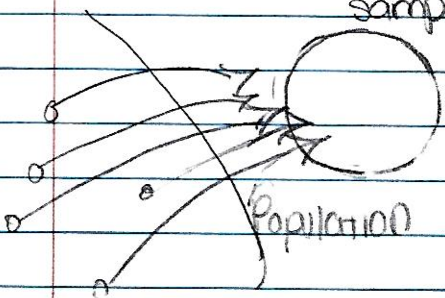
Do you like mushrooms?

$\hat{p} = \dfrac{\# \text{yes's}}{n} = \dfrac{19}{32} = .59$

→ Statistical Inference

Infer the population parameter using the statistic of the data, the $\hat{p}$

What constitutes a sample? → a representative sample
↳ completely random sampling



size $n \ll \infty$

size $N \approx \infty$

$p = \sum_{i=1}^{N} x_i$ unknowable "true" parameter

$X_1, \ldots, X_N \overset{iid}{\sim} \text{Bern}(p)$

We want in the sample size.

$X_1, \ldots, X_n \overset{iid}{\sim} \text{Bern}(\hat{p})$

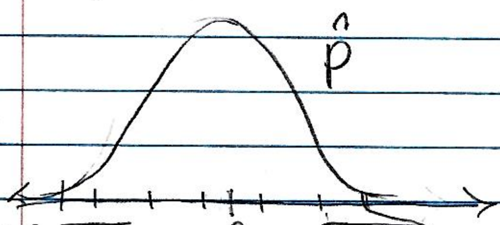
n ← sample size

What type of sample?

You want the sample to be

* completely random selection

*



$p-3\sqrt{\frac{p(1-p)}{n}}$ $\quad p \quad$ $p+\sqrt{\frac{p(1-p)}{n}}$ $\quad p+2\sqrt{\frac{p(1-p)}{n}}$

Goals of Inference

1- Give me best "guess" of p
2- give me a "reasonable" interval of likely values
3- Test "theories" about p.

1) Point Estimation.

$p \approx \hat{p}$

↳ when we don't know p

$\hat{p}$ is the best we can do

2) Confidence interval

Consider the interval

$\left[\hat{p} \pm \sqrt{\frac{p(1-p)}{n}}\right] = \left[\hat{p} - \sqrt{\frac{p(1-p)}{n}}, \hat{p} + \sqrt{\frac{p(1-p)}{n}}\right]$

What is the prob of the interval?

$P\left(p \in \left[\hat{p} \pm \sqrt{\frac{p(1-p)}{n}}\right]\right)$

$\{p, \hat{p}, \hat{P}, P(\ )\}$

$\{[2 \pm 1] = [1, 3]\}$

$$P\left(\hat{p} - \sqrt{\frac{p(1-p)}{n}} \le p \le \hat{p} + \sqrt{\frac{p(1-p)}{n}}\right)$$

$$= P\left(-\sqrt{\frac{p(1-p)}{n}} \le p - \hat{p} \le \sqrt{\frac{p(1-p)}{n}}\right)$$

$$= P\left(-1 \le \frac{p-\hat{p}}{\sqrt{\frac{p(1-p)}{n}}} \le 1\right)$$

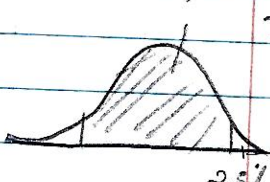
$$P(-1 \le -Z \le 1) \rightarrow = P(1 \ge Z \ge -1)$$

$$= P(Z \in [-1, 1]) = .68$$

$$\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}$$

$$Z_{\frac{\alpha}{2}} := F_Z^{-1}\left(1-\frac{\alpha}{2}\right) \Rightarrow 1-\frac{\alpha}{2} = \int_{-\infty}^{\infty} f_Z(x)\,dx$$

If $\alpha = 10\%$, $\frac{\alpha}{2} = 5\%$, then $1 - \frac{\alpha}{2} = 95\%$.



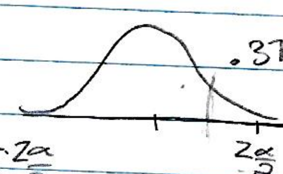
$$P\left(p \in \left[\hat{p} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n}}\right]\right)$$

$$Z_{\frac{\alpha}{2}} := F_Z^{-1}\left(1-\frac{\alpha}{2}\right)$$

$$1-\frac{\alpha}{2} = \int_{-\infty}^{Z_{\frac{\alpha}{2}}} f_Z(x)\,dx$$

$$1-\frac{5\%}{2} = .975 = \int^{Z_{2.5\%}} f_Z(x)\,dx$$

$$Z_{\frac{5\%}{2}} = Z_{2.5\%}$$



$$\frac{\alpha}{2} = \int_{-\infty}^{-Z_{\frac{\alpha}{2}}} f_Z(x)\,dx = F_Z\left(\frac{-\alpha}{2}\right)$$

$$= P\left(\hat{p} - Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}} \le p \le \left(\hat{p} + Z_{\frac{\alpha}{2}}\sqrt{\frac{p(1-p)}{n}}\right)\right)$$

$$= P\left(Z \in \left[-Z_{\frac{\alpha}{2}}, Z_{\frac{\alpha}{2}}\right]\right) = \left(1-\frac{\alpha}{2}\right) - \frac{\alpha}{2} = \boxed{1-\alpha}$$

$$\int_{Z_{\frac{\alpha}{2}}}^{\infty} f_Z(x)\,dx = 1 - \left(1-\frac{\alpha}{2}\right) = \frac{\alpha}{2} = \int_{-\infty}^{-Z_{\frac{\alpha}{2}}} f_Z(x)\,dx$$

92% $\alpha = 8\%$ $Z_{4\%}$

85% $\alpha = 15\%$ $Z_{7.5\%}$

80% $\alpha = 20\%$ $Z_{10\%}$

— Back to mushroom example.

$$\left[0.48 \pm 2\sqrt{\frac{p(1-p)}{n}}\right]$$