this time: frequentist,

next time: Bayesian inference with $\theta \in (0, 1)$

read: J ch. 1–3, Appendices A, B; G ch 1–2

AMS 206
30 Jan 18

1

take-home test 1 problem 1(B)

population: possible outcomes on a single spin

sample & the observed spins

(red = odd)

1 = red
0 = not red

fairness

IID

0	0
0	00
1	1
0	2
1	3
⋮	⋮
0	36

$N = 38$

20 0s
18 1s

mean $\theta = \frac{18}{38} \doteq 0.474$

cumulative relative frequency of 1s (red)

0	0%
1	50
1	67
0	50
0	40
1	50
1	57
1	63
⋮	

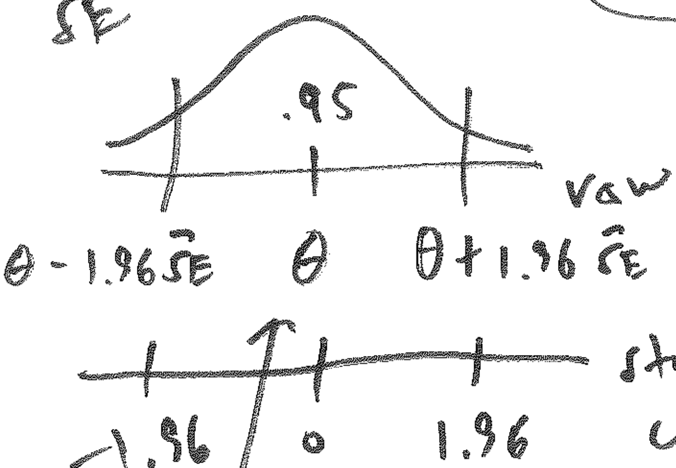
∞

long-run limit $= \theta = \frac{18}{38}$ ✓

weak Law of Large Numbers

$\hat{SE}$ 1.9%

repeated - sampling dist. of $\bar{Y}_n$ Neyman frequentist (F)



random

correct

$$0.95 \doteq P_F\left(\boxed{\theta - 1.96\,\hat{SE} \le \boxed{\bar{Y}} \le \theta + 1.96\,\hat{SE}}\right)$$

fixed unknown constant

Neyman's confidence trick:

$$0.95 \doteq P_F\left(\underline{\bar{Y} - 1.96\,\hat{SE}} \le \theta \le \underline{\bar{Y} + 1.96\,\hat{SE}}\right)$$

So Neyman says: let's use (large n)

$\boxed{\bar{Y} \pm 1.96\,\hat{SE}(\bar{Y})}$ as an (approx.)

95% confidence interval (CI) for θ

here $\bar{Y} = 0.18$

$\widehat{SE}(\bar{Y}) = 0.019$

$\underset{\bar{Y}}{0.18} - 1.96(0.019) \simeq .14$

$0.18 + 1.96(0.019) \simeq .22$

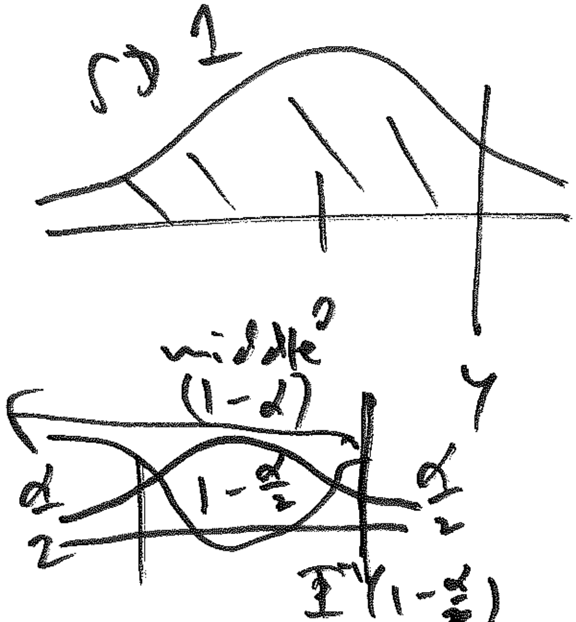
$\bar{Y} = \hat{\theta}$

α = error rate of CI

$100(1-\alpha)\%$ CI:

$$\bar{Y} \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\frac{\bar{Y}(1-\bar{Y})}{n}}$$

$$\hat{\theta} \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}$$



$\Phi(y)$ = standard normal CDF at y

$\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$

↖ quantile

$\Phi(y)$, $y \in \mathbb{R}$, in R is pnorm(y)

$\Phi^{-1}(p)$, $p \in (0,1)$, in R is qnorm(p)

p	$\Phi^{-1}(1-p)$
.01	
.05	. . .
.10	

95% CI in this problem for θ is

$(.142, .218)$

$$P_F(.142 \le \underset{\text{fixed}}{\theta} \le .218) \overset{?}{=} \cancel{0.95}$$

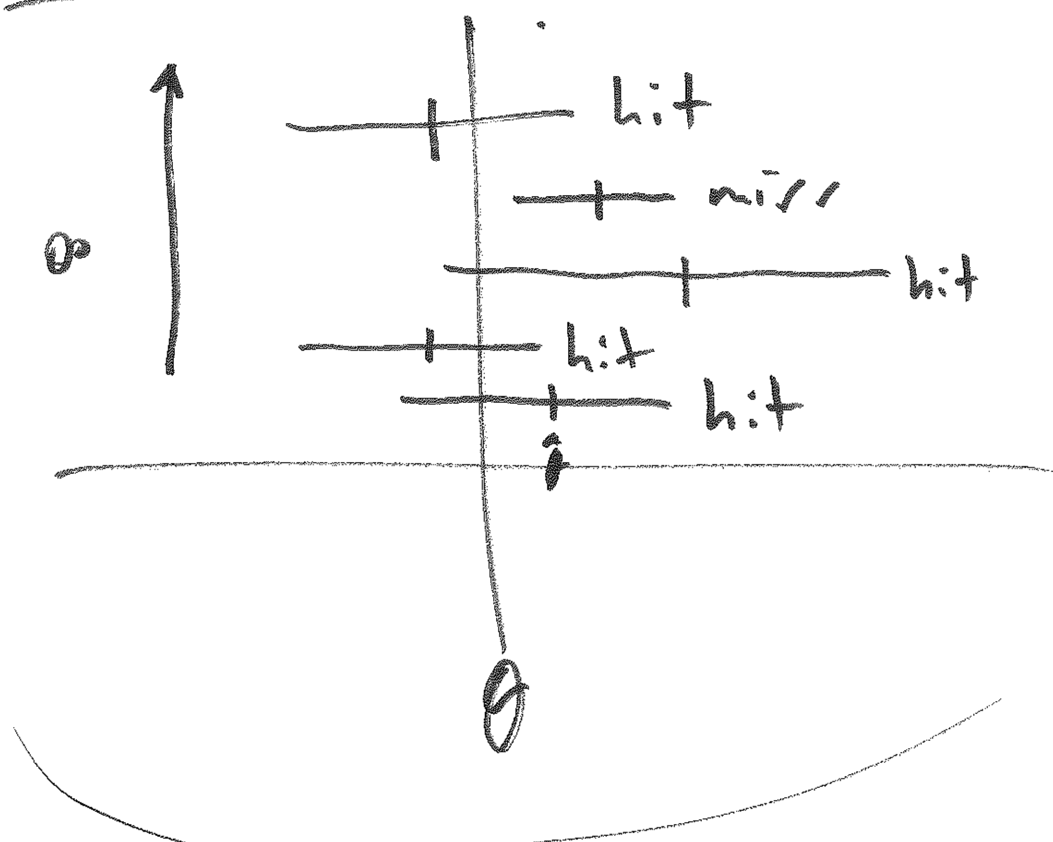
$= \text{undefined}$

confidence

$\neq$ probability

$$P_F\left(\underline{\hat{\theta} - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}} \le \theta \le \underline{\dots (+)}\right) \doteq 95\% \checkmark$$

$$P_F(\underline{0.142} \le \theta \le \underline{0.218}) \ne .95$$



Neyman: about 95% of the CIs will be hits

your confidence is in the

<u>process</u> of CI-building,

not in the outcomes

---

inference about population quantities (parameters) is typically not verifiable

Fisher

$(Y_i \mid \theta) \overset{IID}{\sim} \text{Bernoulli}(\theta)$

discrete $\quad (i = 1, \ldots, n)$

maximum likelihood estimation (MLE)

① write down joint sampling dist. of data random variables:

and

$$\boxed{P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n \mid \theta)}$$

joint

$$\underset{\sim}{y} = (y_1, \ldots, y_n)$$

$$\underset{\sim}{Y} = (Y_1, \ldots, Y_n)$$

$$P(\underset{\sim}{Y} = \underset{\sim}{y} \mid \theta)$$

marginals

$$\overset{IID}{=} \boxed{P(Y_1 = y_1 \mid \theta)} \cdot \ldots \cdot \boxed{P(Y_n = y_n \mid \theta)}$$

$$= \prod_{i=1}^{n} P(Y_i = y_i \mid \theta)$$

$$P(Y_i = y_i \mid \theta) = \begin{cases} \theta & \text{if } y_i = 1 \\ 1-\theta & y_i = 0 \end{cases}$$

$0, 1$

$$= \theta^{y_i} \cdot (1-\theta)^{1-y_i}$$

---

so $P(Y_1 = y_1, \ldots, Y_n = y_n \mid \theta)$

$$= \prod_{i=1}^{n} P(Y_i = y_i \mid \theta)$$

$$= \prod_{i=1}^{n} \theta^{y_i} (1-\theta)^{1-y_i}$$

$$= \theta^{y_1}(1-\theta)^{1-y_1} \cdot \theta^{y_i}(1-\theta)^{1-y_2} \cdot$$

$$\ldots \quad \theta^{y_n}(1-\theta)^{1-y_n}$$

$$= \theta^{y_1+\cdots+y_n}(1-\theta)^{n-(y_1+\cdots+y_n)}$$

let's let $s = s_n = \sum_{i=1}^{n} y_i$ then

$$P(\underset{\sim}{Y} = y \mid \theta) = \theta^s (1-\theta)^{n-s}$$

joint sampling dist ← f'n of y for fixed θ

(2) define likelihood function

a positive ~~constant multiple of~~

as joint sampling dist., but

thought of as a function

of θ for fixed y:

$$\underbrace{\ell(\theta \mid y)}_{\text{likelihood f'n for } \theta \text{ given } y} = c \cdot P(\underset{\sim}{Y} = y \mid \theta)$$

↑ arbitrary positive constant

$$= c \cdot \theta^s (1-\theta)^{n-s}$$

def. if $\ell(\theta | y)$ depends on y only through a summary of y (e.g., $s = \sum_{i=1}^{n} y_i$), that summary is said to be (a sufficient statistic) for θ with this likelihood

sufficient

- y vector of length n

s —— 1

$n = 400$

$s = 72$

$$\theta^{72} (1-\theta)^{400-72}$$