

Goodness of fitEx: Coin thrown until heads. ( $n=250$ )

# tosses 1 2 3 4 5 6 7 8

freq 130 60 34 12 9 1 3 1

6, 7, or 8, 5 count.

Does Data fit a geo ( $P=\frac{1}{2}$ ) dist at level  $\alpha=0.05$ ? $H_0: X \sim \text{Geom}(\frac{1}{2})$  vs  $H_1: X \not\sim \text{Geom}(\frac{1}{2})$ under  $H_0$ Expected  
Freq:

125, 62.5, 31.15, 15.575, 7.8125, 3.9, 1.82, 0.91, ...

$$\chi^2 = \frac{(130-125)^2}{125} + \frac{(60-62.5)^2}{62.5} + \dots \quad (\text{group last 3 terms})$$

New table:

X	1	2	3	4	5	6+
freq	130	60	34	12	9	5
EF	125	62.5	31.15	15.575	7.8125	7.8125

$$\chi^2 \sim \chi^2_{6-0-1} = \chi^2_5 \quad \text{so CR is } \chi^2 \geq \chi^2_{0.05,5} = 11.07$$

$$\chi^2 = \frac{(130-125)^2}{125} + \frac{(60-62.5)^2}{62.5} + \dots + \frac{(5-7.8125)^2}{7.8125} = 2.577$$

so fail to reject  $H_0$ , conclude that  $X \sim \text{Geom}(\frac{1}{2})$  is a good explanation / fit for the data at level  $\alpha=0.05$ .

What if we don't have a specific  $\theta$  or  $p$  or whatever?Question  $X \sim \text{Geom}(p)$ ? ( $p$  not specified).

First, estimate  $p$  using data. now  $t=1$  so  $X^2 \sim \chi^2_4$  in this example.

use  $P = \frac{1}{\bar{X}} = 0.5028$ .

So table:

X	1	2	3	4	5	$\geq 6$
freq	130	60	34	12	9	5
exp freq	130.2	62.5	29.9	14.33	6.87	6.31

NOTE: only combine columns when expected counts are  $< 5$ .

$$\chi^2 = 1.57 \leq \chi^2_{0.05, 4} = 9.488 \Rightarrow \text{fail to reject } H_0.$$

★ Goodness of fit: POISSON for EXAM