April 23rd, 2025

Goodness of Fit

Example:

- Lets just say you want to write a standardized exam where every answer (a, b, c, ...) have an equal chance of being the answer
- Extreme example: 2 Questions with 2 possible answers
  - Q1, Answer: B. So Q2, Answer: A, then its even ← This is not a good way to construct the test, in the second question, only A can be chosen.
  - New way: Flip a coin for each question, heads - A, tails - B. This is a random process. For even odds you want the probability of A, $P_A = 0.5$, and $P_B = 0.5$.
  - Now lets say you made a machine to randomly assign an answer to each question

A	B
47	53
45	55

→ How do we check if this is random?

- Real Example: 400 questions with 5 possible answers

	A	B	C	D	E	Total
Expected	80	80	80	80	80	400
Observed	74	90	76	87	73	400

$H_0: \begin{cases} P_A = 0.2 \\ P_B = 0.2 \\ \vdots \\ P_E = 0.2 \end{cases}$

The t-statistic

$$t = \frac{\bar{x} - \mu_0}{\hat{\sigma}/\sqrt{n}}$$

- Test statistic
- Relates observed to null (expected)
- When H_0 true, follows known distribution

χ^2 test

$$\chi^2 = \sum_{i=1}^{N} \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

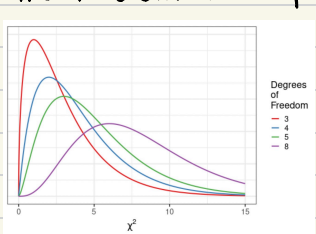
- Testing "goodness of fit"
- "Chi-squared" ("Kai")

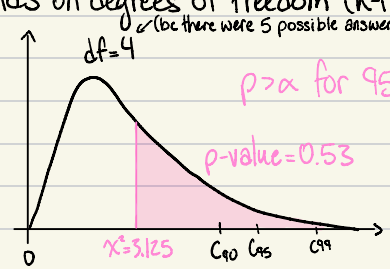
DF: The number of observations that are independent. If you have 3 possibilities on a multiple choice test: a, b, c. And 10 questions and 4 questions have answer (a) and 3 have (b), you know 3 must have answer (c) to get to 10 answers. Then only (a) and (b) are independent so df = 2.

- χ^2 on our example: $\chi^2 = \frac{(80-74)^2}{80} + \frac{(80-90)^2}{80} + \frac{(80-76)^2}{80} + \frac{(80-87)^2}{80} + \frac{(80-73)^2}{80}$

→ $= 3.125$

- What does this number mean?
  - The χ^2 distribution depends on degrees of freedom (k-1) (bc there were 5 possible answers)





$p > \alpha$ for 95% confidence so fail to reject

New Example: Looking for racial bias in the jury pool of Alameda County
Below is the racial and ethnic composition of 1453 individuals in the jury pool along with the census proportions:

Race Ethnicity	White	Black	Hispanic	Asian	Other	Total
Jury Size	780	117	114	384	58	1453
Census Percentage	54%	18%	12%	15%	1%	100%
Expected	784.62	261.54	174.36	217.95	14.53	1453

Total·percentage 1453·0.54

$$\chi^2 = 357 = \frac{(784.62-780)^2}{784.62} + \frac{(261.54-117)^2}{261.54} + \dots + \frac{(14.53-58)^2}{14.53}$$

$C_{95} = 9.49$ (for $df = 4$)

$\chi^2 > C_{95}$ so we will reject the null hypothesis. That is, we conclude that the jury pool is not representative of the population. The χ^2 test however does not tell us anything about which groups are over or underrepresented.