V33 - Hypothesis Testing - Part 9

- Goodness of Fit Test

Course: Statistical Testing & Regression
Dr. Renee Clark
Swanson School of Engineering
Industrial Engineering
University of Pittsburgh

Goodness of Fit Test

- Thus far, tested hypotheses about population parameters:
  μ, p, σ^2, $\mu_1 - \mu_2$, σ_1^2/σ_2^2, $p_1 - p_2$
- Next → test hypotheses about data distributions
  - Does *observed* or *sample* distribution follow certain theoretical or hypothesized distribution?
  - Example: does observed data follow a normal or uniform distribution?
- HT known as Goodness of Fit test between 2 distributions:
  - Observed vs. Theoretical distribution
    1 2

Goodness of Fit Test

- Example: tossing a die
  - Outcomes have uniform distribution (assuming die fair → equal prob)
  - Discrete uniform (6 sides)
- H_0: Outcomes have **discrete uniform distribution**
  - each face equally likely to occur
  - probability function for a fair die: $\boxed{f(x) = \frac{1}{6}}$ $x = 1,2,3,4,5,6.$
  - 6 equal probabilities
- If toss die 120 times and die is fair, we **expect** each face to appear 20 times

$\frac{120}{6} = 20$

Outcome	Expected Frequency
1	20
2	20
3	20
4	20
5	20
6	20

Goodness of Fit Test

- However, observed/actual frequencies may not match expected
- GF test explores whether differences due to chance
  - …or whether the distribution is *not* uniform discrete (die not fair)

Outcome	Expected Frequency	Observed Frequency
1	20	20
2	20	22
3	20	17
4	20	18
5	20	19
6	20	24
	/20	/20

Acknowledgement

This material is based upon work partially supported by the National Science Foundation under Grant# 2335802. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

You are free

- to **Share** – to copy, distribute, display and perform the work
- to **Remix** – to make derivative works

Under the following conditions

- **Attribution** — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- **Noncommercial** — You may not use this work for commercial purposes.
- **Share Alike** — If you alter, transform, or build upon work, you may distribute the resulting work only under the same or similar license to this one.

THE END