Decisions and Hypothesis Testing

PUBH 8442: Bayes Decision Theory and Data Analysis

Eric F. Lock
UMN Division of Biostatistics, SPH
elock@umn.edu

2/12/2025

Example: Coke bottles (cont.)

- Recall:
  - Coke bottles are filled with calibration $\text{Normal}(12, 0.01)$
  - Given machine with calibration μ, bottles filled with $\text{Normal}(\mu, 0.05)$
  - For $n = 5$, $p(\mu \mid \mathbf{y}) = \text{Normal}(\frac{1}{2}(12 + \bar{y}), 0.005)$
- A machine with calibration μ costs the company

$$\$50000(\mu - 12)^2$$

- The cost to re-calibrate a machine to $\mu = 12$ is \$500
- Decide whether to re-calibrate after sample of $n = 5$

Example: Coke bottles room (cont.)

- Decision theoretic framework:
  - *prior distribution:* $p(\mu) = \text{Normal}(12, 0.01)$, $\mu \in \mathbb{R}$
  - *sampling distribution:* $\{y_i\}_{i=1}^{5} \overset{iid}{\sim} \text{Normal}(\mu, 0.05)$, $\mathbf{y} \in \mathbb{R}^5$
  - *allowable actions:*

$$\mathcal{A} = \{\text{Re-calibrate}(R), \text{Do not re-calibrate}(N)\}$$

  - *loss function:*

$$l(\mu, a) = \begin{cases} 500 & \text{if } a = R \\ 50000(\mu - 12)^2 & \text{if } a = N \end{cases}$$

  - *decision rule:* $d(y) = ?$

Example: Coke bottles room (cont.)

- The posterior risk for not recalibrating is

$$\rho(p_\theta, N) = 250 + 12500(\bar{y} - 12)^2$$

Example: Coke bottles room (cont.)

- $\rho(p_\theta, N) < \rho(p_\theta, R) = 500$ when $\bar{y} \in (11.86, 12.14)$

- The Bayes decision rule is

$$d(\mathbf{y}) = \begin{cases} \text{Do not recalibrate} & \text{if } \bar{y} \in (11.86, 12.14) \\ \text{Recalibrate} & \text{if } \bar{y} \notin (11.86, 12.14) \end{cases}$$

Frequentist hypothesis testing

- Two possible hypotheses involving θ: the null (H_0) and alternative (H_a)
- Observe data $\mathbf{y}$, and assume $\mathbf{Y}$ and $\mathbf{y}$ are iid given θ
- The *p-value* is the probability that $\mathbf{Y}$ are more "extreme" than $\mathbf{y}$, under H_0:

$$\text{p-value} = P(\mathbf{Y} \text{ more "extreme" than } \mathbf{y} \mid H_0)$$

  - $\mathbf{Y}$ represents results that "could have" occurred under H_0
  - "Extreme" depends on context
  - Probability always computed under H_0, never H_a
  - H_0 is usually more specific: e.g., $H_0 : \theta = \theta_0, H_a : \theta \neq \theta_0$
  - Small p-value is evidence against H_0

Frequentist hypothesis testing

- p-value requires considering probability of other possible outcomes, not just the observed outcome
- Therefore, it violates the *Likelihood Principle*:
  - In making inferences or decisions about θ after $\mathbf{y}$ is observed, all relevant experimental information is contained in the likelihood function for the observed $\mathbf{y}$.
- Two experiments can give data with different p-values but equal likelihood under H_0

Example: tipping pennies

- Claim: If you stand a penny on its side and let it fall, it will land heads more often than tails.
  - Let $\theta =$ probability the penny lands heads
  - $H_0 : \theta = 1/2$
  - $H_a : \theta > 1/2$
- Experiment 1: tip the coin until the penny lands tails
  - Consider X : number of heads before first tail
- Experiment 2: tip the coin 6 times
  - Consider Y : number of times penny falls heads

Example: tipping pennies

- Given θ,
  - $X \sim \text{NegativeBinomial}(1, \theta)$
  - $Y \sim \text{Binomial}(6, \theta)$
- Assume for both experiments we observe 5 heads, then one tail.
- Gives identical likelihoods:

$$P(X = 5 \mid \theta) \propto P(Y = 5 \mid \theta) \propto \theta^5(1 - \theta)$$

Example: tipping pennies (cont.)

- For experiment 1, p-value is $P(X \geq 5 \mid H_0) = 0.031$

- For experiment 2, p-value is $P(Y \geq 5 \mid H_0) = 0.110$

Bayesian hypothesis testing

- Could compute $P(H_0 \mid \mathbf{y})$ directly:

$$P(H_0 \mid \mathbf{y}) = \frac{P(\mathbf{y} \mid H_0)P(H_0)}{P(\mathbf{y} \mid H_0)P(H_0) + P(\mathbf{y} \mid H_a)P(H_a)}$$

  - Must compute $P(\mathbf{y} \mid H_a)$ – requires prior for H_a
  - Must specify $P(H_0)$, our prior probability of H_0

- $P(H_0 \mid \mathbf{y}) + P(H_a \mid \mathbf{y}) = 1$
- Symmetric: can give evidence for or against H_0
  - P-values can only give evidence against (never "accept H_0")

Example: tipping pennies

- Possible Bayesian framework:
  - Let $\theta =$ probability the penny lands heads
  - $H_0 : \theta = 1/2$
  - $H_a : \theta \sim \text{Uniform}(0.5, 1)$
  - $P(H_0) = 0.5$
- Our prior for θ is

$$\theta \sim \begin{cases} 1/2 & \text{with probability } 1/2 \\ p(\theta) = 2 \text{ for } \theta \in [0.5, 1] & \text{with probability } 1/2 \end{cases}$$

Example: tipping pennies

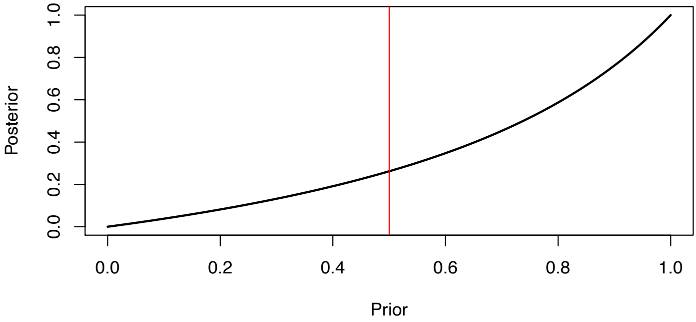
- For experiment 1,
  - $P(x = 5 \mid H_0) = (1/2)^6$
  - $P(x = 5 \mid H_a) \approx 0.044$
- So, $P(H_0 \mid x = 5) \approx 0.267$

Example: tipping pennies

- For experiment 2,
  - $P(y = 5 \mid H_0) = 6(1/2)^6$
  - $P(y = 5 \mid H_a) \approx 0.264$
- So, $P(H_0 \mid y = 5) \approx 0.267$

Example: tipping pennies

- Posterior $P(H_0 \mid y = 5)$ as function of prior $P(H_0)$:



-

Code: http://www.ericfrazerlock.com/Decisions_and_Hypothesis_Testing_Rcode1.r

Bayesian hypothesis testing

- In general, if $P(\mathbf{y} \mid \theta) \propto P(\mathbf{x} \mid \theta)$, then

$$P(H_0 \mid \mathbf{y}) = P(H_0 \mid \mathbf{x})$$

- Analogous result holds for continuous x,y with $p(\mathbf{y} \mid \theta) \propto p(\mathbf{x} \mid \theta)$

- Thus, Bayesian hypothesis testing satisfies the likelihood principle