Hypothesis Testing

Prof. Pradeep Ravikumar
pradeepr@cs.cmu.edu

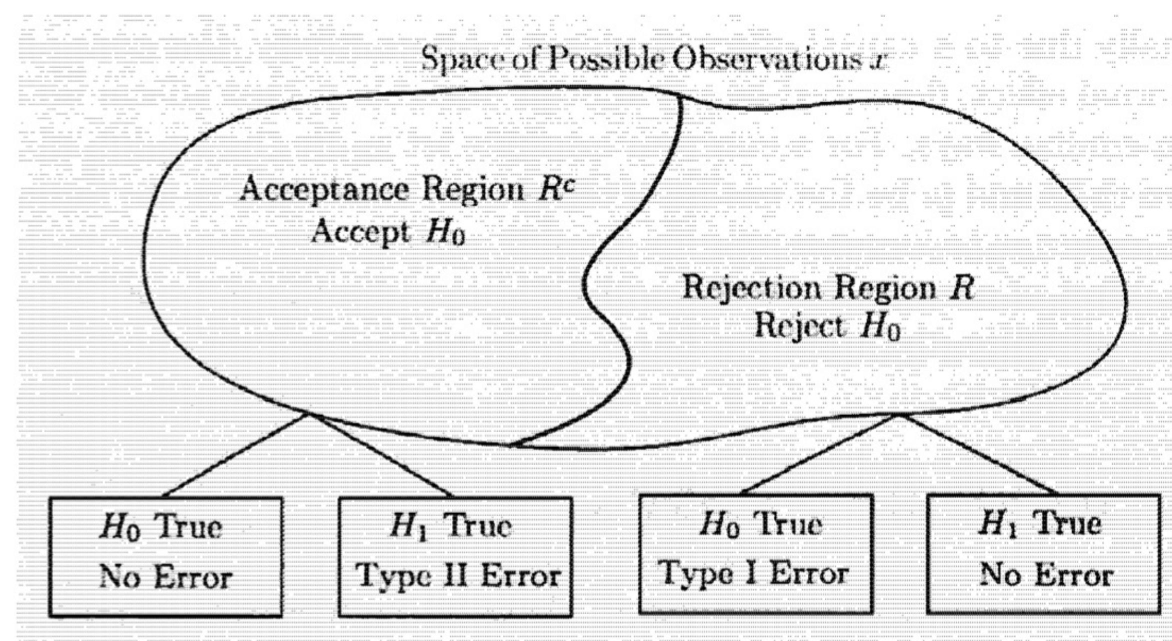
Recall: Binary Hypothesis Testing

- Partition the space of possible data vectors
  Rejection region R:
  reject H_0 iff data $\in R$

- Types of errors:
  - **Type I** (**false rejection**, false alarm):
    H_0 true, but rejected

$$\alpha(R) = \mathbf{P}(X \in R; H_0)$$

  - **Type II** (**false acceptance**, missed detection):
    H_0 false, but accepted



Likelihood Ratio Test (LRT)

- Bayesian case (MAP rule): choose H_1 if:
  $\mathbf{P}(H_1 \mid X = x) > \mathbf{P}(H_0 \mid X = x)$
  or
  $\dfrac{\mathbf{P}(X = x \mid H_1)\mathbf{P}(H_1)}{\mathbf{P}(X = x)} > \dfrac{\mathbf{P}(X = x \mid H_0)\mathbf{P}(H_0)}{\mathbf{P}(X = x)}$
  or
  $$\frac{\mathbf{P}(X = x \mid H_1)}{\mathbf{P}(X = x \mid H_0)} > \frac{\mathbf{P}(H_1)}{\mathbf{P}(H_0)}$$
  (likelihood ratio test)

Likelihood Ratio Test (LRT)

- Bayesian case (MAP rule): choose H_1 if:
  $\mathbf{P}(H_1 \mid X = x) > \mathbf{P}(H_0 \mid X = x)$
  or
  $$\frac{\mathbf{P}(X = x \mid H_1)\mathbf{P}(H_1)}{\mathbf{P}(X = x)} > \frac{\mathbf{P}(X = x \mid H_0)\mathbf{P}(H_0)}{\mathbf{P}(X = x)}$$
  or
  $$\frac{\mathbf{P}(X = x \mid H_1)}{\mathbf{P}(X = x \mid H_0)} > \frac{\mathbf{P}(H_1)}{\mathbf{P}(H_0)}$$
  (likelihood ratio test)

- Nonbayesian version: choose H_1 if
  $$\frac{\mathbf{P}(X = x; H_1)}{\mathbf{P}(X = x; H_0)} > \xi \quad \text{(discrete case)}$$
  $$\frac{f_X(x; H_1)}{f_X(x; H_0)} > \xi \qquad \text{(continuous case)}$$

Likelihood Ratio Test (LRT)

- Bayesian case (MAP rule): choose H_1 if:
  $\mathbf{P}(H_1 \mid X = x) > \mathbf{P}(H_0 \mid X = x)$
  or
  $$\frac{\mathbf{P}(X = x \mid H_1)\mathbf{P}(H_1)}{\mathbf{P}(X = x)} > \frac{\mathbf{P}(X = x \mid H_0)\mathbf{P}(H_0)}{\mathbf{P}(X = x)}$$
  or
  $$\frac{\mathbf{P}(X = x \mid H_1)}{\mathbf{P}(X = x \mid H_0)} > \frac{\mathbf{P}(H_1)}{\mathbf{P}(H_0)}$$
  (likelihood ratio test)

- Nonbayesian version: choose H_1 if
  $$\frac{\mathbf{P}(X = x; H_1)}{\mathbf{P}(X = x; H_0)} > \xi \quad \text{(discrete case)}$$
  $$\frac{f_X(x; H_1)}{f_X(x; H_0)} > \xi \qquad \text{(continuous case)}$$

- threshold ξ trades off the two types of error
  - choose ξ so that $\mathbf{P}(\text{reject } H_0; H_0) = \alpha$ (e.g., $\alpha = 0.05$)

Example I

We have a six-sided die that we want to test for fairness, and we formulate two hypotheses for the probabilities of the six faces:

$$H_0 \text{ (fair die):} \qquad p_X(x; H_0) = \frac{1}{6}, \qquad x = 1, \ldots, 6,$$

$$H_1 \text{ (loaded die):} \qquad p_X(x; H_1) = \begin{cases} \frac{1}{4}, & \text{if } x = 1, 2, \\ \frac{1}{8}, & \text{if } x = 3, 4, 5, 6. \end{cases}$$

Example I

We have a six-sided die that we want to test for fairness, and we formulate two hypotheses for the probabilities of the six faces:

$$H_0 \text{ (fair die):} \qquad p_X(x; H_0) = \frac{1}{6}, \qquad x = 1, \ldots, 6,$$

$$H_1 \text{ (loaded die):} \qquad p_X(x; H_1) = \begin{cases} \frac{1}{4}, & \text{if } x = 1, 2, \\ \frac{1}{8}, & \text{if } x = 3, 4, 5, 6. \end{cases}$$

The likelihood ratio for a single roll x of the die is

$$L(x) = \begin{cases} \frac{1/4}{1/6} = \frac{3}{2}, & \text{if } x = 1, 2, \\ \frac{1/8}{1/6} = \frac{3}{4}, & \text{if } x = 3, 4, 5, 6. \end{cases}$$

Example I

We have a six-sided die that we want to test for fairness, and we formulate two hypotheses for the probabilities of the six faces:

The likelihood ratio for a single roll x of the die is

$$L(x) = \begin{cases} \dfrac{1/4}{1/6} = \dfrac{3}{2}, & \text{if } x = 1, 2, \\ \dfrac{1/8}{1/6} = \dfrac{3}{4}, & \text{if } x = 3, 4, 5, 6. \end{cases}$$

$$\xi < \frac{3}{4}: \qquad \text{reject } H_0 \text{ for all } x;$$

Example I

We have a six-sided die that we want to test for fairness, and we formulate two hypotheses for the probabilities of the six faces:

The likelihood ratio for a single roll x of the die is

$$L(x) = \begin{cases} \dfrac{1/4}{1/6} = \dfrac{3}{2}, & \text{if } x = 1, 2, \\ \dfrac{1/8}{1/6} = \dfrac{3}{4}. & \text{if } x = 3, 4, 5, 6. \end{cases}$$

$$\begin{aligned} \xi < \frac{3}{4}: &\qquad \text{reject } H_0 \text{ for all } x; \\ \frac{3}{4} < \xi < \frac{3}{2}: &\qquad \text{accept } H_0 \text{ if } x = 3, 4, 5, 6; \ \text{reject } H_0 \text{ if } x = 1, 2; \end{aligned}$$

Example I

We have a six-sided die that we want to test for fairness, and we formulate two hypotheses for the probabilities of the six faces:

The likelihood ratio for a single roll x of the die is

$$L(x) = \begin{cases} \frac{1/4}{1/6} = \frac{3}{2}, & \text{if } x = 1, 2, \\ \frac{1/8}{1/6} = \frac{3}{4}, & \text{if } x = 3, 4, 5, 6. \end{cases}$$

$$\begin{aligned} \xi < \frac{3}{4}: & \quad \text{reject } H_0 \text{ for all } x; \\ \frac{3}{4} < \xi < \frac{3}{2}: & \quad \text{accept } H_0 \text{ if } x = 3, 4, 5, 6; \text{ reject } H_0 \text{ if } x = 1, 2; \\ \frac{3}{2} < \xi: & \quad \text{accept } H_0 \text{ for all } x. \end{aligned}$$

Example I

We have a six-sided die that we want to test for fairness, and we formulate two hypotheses for the probabilities of the six faces:

$$\begin{aligned} \xi < \frac{3}{4}: &\quad \text{reject } H_0 \text{ for all } x; \\ \frac{3}{4} < \xi < \frac{3}{2}: &\quad \text{accept } H_0 \text{ if } x = 3, 4, 5, 6; \ \text{reject } H_0 \text{ if } x = 1, 2; \\ \frac{3}{2} < \xi: &\quad \text{accept } H_0 \text{ for all } x. \end{aligned}$$

probability of false rejection $\mathbf{P}(\text{Reject } H_0; H_0)$

$$\alpha(\xi) = \begin{cases} 1, & \text{if } \xi < \frac{3}{4}, \\ \mathbf{P}(X = 1, 2; H_0) = \frac{1}{3}, & \text{if } \frac{3}{4} < \xi < \frac{3}{2}, \\ 0, & \text{if } \frac{3}{2} < \xi. \end{cases}$$

Example I

We have a six-sided die that we want to test for fairness, and we formulate two hypotheses for the probabilities of the six faces:

$$
\begin{aligned}
\xi < \frac{3}{4}: &\quad \text{reject } H_0 \text{ for all } x; \\
\frac{3}{4} < \xi < \frac{3}{2}: &\quad \text{accept } H_0 \text{ if } x = 3, 4, 5, 6; \text{ reject } H_0 \text{ if } x = 1, 2; \\
\frac{3}{2} < \xi: &\quad \text{accept } H_0 \text{ for all } x.
\end{aligned}
$$

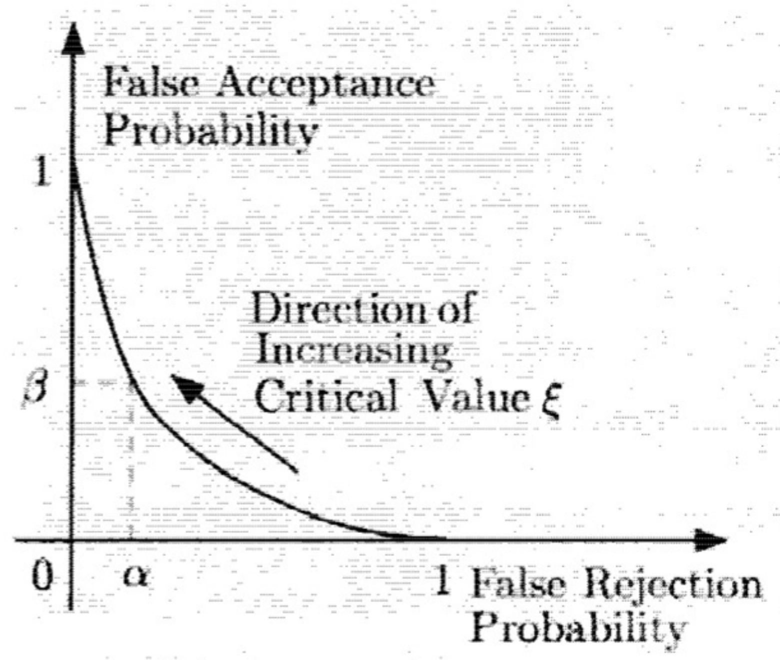
probability of false rejection $\mathbf{P}(\text{Reject } H_0; H_0)$

$$
\alpha(\xi) = \begin{cases} 1, & \text{if } \xi < \frac{3}{4}, \\ \mathbf{P}(X = 1, 2; H_0) = \frac{1}{3}, & \text{if } \frac{3}{4} < \xi < \frac{3}{2}, \\ 0, & \text{if } \frac{3}{2} < \xi. \end{cases}
$$

probability of false acceptance $\mathbf{P}(\text{Accept } H_0; H_1)$

$$
\beta(\xi) = \begin{cases} 0, & \text{if } \xi < \frac{3}{4}, \\ \mathbf{P}(X = 3, 4, 5, 6; H_1) = \frac{1}{2}, & \text{if } \frac{3}{4} < \xi < \frac{3}{2}, \\ 1, & \text{if } \frac{3}{2} < \xi. \end{cases}
$$

Example I: Tradeoff



probability of false rejection $\mathbf{P}(\text{Reject } H_0; H_0)$

$$\alpha(\xi) = \begin{cases} 1, & \text{if } \xi < \frac{3}{4}, \\ \mathbf{P}(X = 1, 2; H_0) = \frac{1}{3}, & \text{if } \frac{3}{4} < \xi < \frac{3}{2}, \\ 0, & \text{if } \frac{3}{2} < \xi. \end{cases}$$

probability of false acceptance $\mathbf{P}(\text{Accept } H_0; H_1)$

$$\beta(\xi) = \begin{cases} 0, & \text{if } \xi < \frac{3}{4}, \\ \mathbf{P}(X = 3, 4, 5, 6; H_1) = \frac{1}{2}, & \text{if } \frac{3}{4} < \xi < \frac{3}{2}, \\ 1, & \text{if } \frac{3}{2} < \xi. \end{cases}$$

Summary: Likelihood Ratio Test

- Start with a target value α for the false rejection probability.
- Choose a value for ξ such that the false rejection probability is equal to α:
$$\mathbf{P}\big(L(X) > \xi; H_0\big) = \alpha.$$
- Once the value x of X is observed, reject H_0 if $L(x) > \xi$.

Example II

A surveillance camera checks a certain area and record a signal $X = W$, or $X = 1 + W$, depending on whether an intruder is absent or present (hypotheses H_0 or H_1 respectively).

Assume that $W \sim N(0,v)$, for some known $v > 0$. Write out the Likelihood Ratio Test for testing the presence of an intruder (i.e. hypothesis H_0 or H_1), given the camera signal x.

Example II

A surveillance camera checks a certain area and record a signal X = W, or X = 1 + W, depending on whether an intruder is absent or present (hypotheses H_0 or H_1 respectively).

Assume that W ~ N(0,v), for some known v > 0. Write out the Likelihood Ratio Test for testing the presence of an intruder (i.e. hypothesis H_0 or H_1), given the camera signal x.

$$f_X(x; H_0) = \frac{1}{\sqrt{2\pi v}} \exp\left\{-\frac{x^2}{2v}\right\}, \qquad f_X(x; H_1) = \frac{1}{\sqrt{2\pi v}} \exp\left\{-\frac{(x-1)^2}{2v}\right\},$$

Likelihood Function:

$$L(x) = \frac{f_X(x; H_1)}{f_X(x; H_0)} = \exp\left\{\frac{x^2 - (x-1)^2}{2v}\right\} = \exp\left\{\frac{2x-1}{2v}\right\}.$$

Example II

A surveillance camera checks a certain area and record a signal X = W, or X = 1 + W, depending on whether an intruder is absent or present (hypotheses H_0 or H_1 respectively).

Assume that W ~ N(0,v), for some known v > 0. Write out the Likelihood Ratio Test for testing the presence of an intruder (i.e. hypothesis H_0 or H_1), given the camera signal x.

Likelihood Ratio:

$$L(x) = \frac{f_X(x; H_1)}{f_X(x; H_0)} = \exp\left\{\frac{x^2 - (x-1)^2}{2v}\right\} = \exp\left\{\frac{2x-1}{2v}\right\}.$$

LRT Test:

$$L(x) > \xi \;\equiv\; x > v \log \xi + \frac{1}{2}$$

Example II

A surveillance camera checks a certain area and record a signal X = W, or X = 1 + W, depending on whether an intruder is absent or present (hypotheses H_0 or H_1 respectively).

Assume that W ~ N(0,v), for some known v > 0. Write out the Likelihood Ratio Test for testing the presence of an intruder (i.e. hypothesis H_0 or H_1), given the camera signal x.

Likelihood Ratio:

$$L(x) = \frac{f_X(x; H_1)}{f_X(x; H_0)} = \exp\left\{\frac{x^2 - (x-1)^2}{2v}\right\} = \exp\left\{\frac{2x-1}{2v}\right\}.$$

LRT Test:

$$L(x) > \xi \;\equiv\; x > v \log \xi + \frac{1}{2}$$

Rejection Region:

$$R = \{x \mid x > \gamma\} \qquad \gamma = v \log \xi + \frac{1}{2};$$

Example II

A surveillance camera checks a certain area and record a signal X = W, or X = 1 + W, depending on whether an intruder is absent or present (hypotheses H_0 or H_1 respectively).

Assume that W ~ N(0,v), for some known v > 0. Write out the Likelihood Ratio Test for testing the presence of an intruder (i.e. hypothesis H_0 or H_1), given the camera signal x.

LRT Test:

$$L(x) > \xi \;\equiv\; x > v \log \xi + \frac{1}{2}$$

Rejection Region:

$$R = \{x \mid x > \gamma\} \qquad \gamma = v \log \xi + \frac{1}{2};$$

$$\alpha = \mathbf{P}(X > \gamma; H_0) = \mathbf{P}(W > \gamma),$$

so that for a specified value of α, we can obtain the value of γ from a table/calculator

Example II

A surveillance camera checks a certain area and record a signal X = W, or X = 1 + W, depending on whether an intruder is absent or present (hypotheses H_0 or H_1 respectively).

Assume that W ~ N(0,v), for some known v > 0. Write out the Likelihood Ratio Test for testing the presence of an intruder (i.e. hypothesis H_0 or H_1), given the camera signal x.

LRT Test:

$$L(x) > \xi \;\equiv\; x > v \log \xi + \frac{1}{2}$$

Rejection Region:

$$R = \{x \,|\, x > \gamma\} \qquad \gamma = v \log \xi + \frac{1}{2};$$

False Rejection Probability:

$$\alpha = \mathbf{P}(X > \gamma; H_0) = \mathbf{P}(W > \gamma),$$

so that for a specified value of α, we can obtain the value of γ from a table/calculator

False Acceptance Probability:

$$\beta = \mathbf{P}\big(X \leq \gamma; H_1\big) = \mathbf{P}\,(1 + W \leq \gamma) = \mathbf{P}\,(W \leq \gamma - 1),$$

Example III

A Discrete Example. Consider $n = 25$ independent tosses of a coin. Under hypothesis H_0 (respectively, H_1), the probability of a head at each toss is equal to $\theta_0 = 1/2$ (respectively, $\theta_1 = 2/3$). Let X be the number of heads observed. If we set the false rejection probability to 0.1, what is the rejection region associated with the LRT?

Example III

A Discrete Example. Consider $n = 25$ independent tosses of a coin. Under hypothesis H_0 (respectively, H_1), the probability of a head at each toss is equal to $\theta_0 = 1/2$ (respectively, $\theta_1 = 2/3$). Let X be the number of heads observed. If we set the false rejection probability to 0.1, what is the rejection region associated with the LRT?

We observe that when $X = k$, the likelihood ratio is of the form

$$L(k) = \frac{\binom{n}{k}\theta_1^k(1-\theta_1)^{n-k}}{\binom{n}{k}\theta_0^k(1-\theta_0)^{n-k}} = \left(\frac{\theta_1}{\theta_0}\cdot\frac{1-\theta_0}{1-\theta_1}\right)^k \cdot \left(\frac{1-\theta_1}{1-\theta_0}\right)^n = 2^k\left(\frac{2}{3}\right)^{25}.$$

Example III

A Discrete Example. Consider $n = 25$ independent tosses of a coin. Under hypothesis H_0 (respectively, H_1), the probability of a head at each toss is equal to $\theta_0 = 1/2$ (respectively, $\theta_1 = 2/3$). Let X be the number of heads observed. If we set the false rejection probability to 0.1, what is the rejection region associated with the LRT?

We observe that when $X = k$, the likelihood ratio is of the form

$$L(k) = \frac{\binom{n}{k}\theta_1^k(1-\theta_1)^{n-k}}{\binom{n}{k}\theta_0^k(1-\theta_0)^{n-k}} = \left(\frac{\theta_1}{\theta_0}\cdot\frac{1-\theta_0}{1-\theta_1}\right)^k \cdot \left(\frac{1-\theta_1}{1-\theta_0}\right)^n = 2^k\left(\frac{2}{3}\right)^{25}.$$

LRT Test: reject H_0 if $X > \gamma$.

Example III

A Discrete Example. Consider $n = 25$ independent tosses of a coin. Under hypothesis H_0 (respectively, H_1), the probability of a head at each toss is equal to $\theta_0 = 1/2$ (respectively, $\theta_1 = 2/3$). Let X be the number of heads observed. If we set the false rejection probability to 0.1, what is the rejection region associated with the LRT?

We observe that when $X = k$, the likelihood ratio is of the form

$$L(k) = \frac{\binom{n}{k}\theta_1^k(1-\theta_1)^{n-k}}{\binom{n}{k}\theta_0^k(1-\theta_0)^{n-k}} = \left(\frac{\theta_1}{\theta_0}\cdot\frac{1-\theta_0}{1-\theta_1}\right)^k \cdot \left(\frac{1-\theta_1}{1-\theta_0}\right)^n = 2^k\left(\frac{2}{3}\right)^{25}.$$

LRT Test: reject H_0 if $X > \gamma$.

Want γ so that false rejection prob. is at most 0.1!

Example III

A Discrete Example. Consider $n = 25$ independent tosses of a coin. Under hypothesis H_0 (respectively, H_1), the probability of a head at each toss is equal to $\theta_0 = 1/2$ (respectively, $\theta_1 = 2/3$). Let X be the number of heads observed. If we set the false rejection probability to 0.1, what is the rejection region associated with the LRT?

LRT Test: **reject H_0 if $X > \gamma$.**

Want γ so that false rejection prob. is at most 0.1!

$\mathbf{P}(X > \gamma; H_0) \leq 0.1$, or

$$\sum_{i=\gamma+1}^{25} \binom{25}{i} 2^{-25} \leq 0.1.$$

Numerical evaluations yield γ = 16

Example IV

Problem : Artemisia moves to a new house and she is "fifty-percent sure" that the phone number is 2537267. To verify this. she uses the house phone to dial 2537267. she obtains a busy signal. and concludes that this is indeed the correct number. Assuming that the probability of a typical seven-digit phone number being busy at any given time is 1%. what is the probability that Artemisia's conclusion was correct?

H_0 : the phone number is 2537267,

H_1 : the phone number is not 2537267,

$$\mathbf{P}(H_0) = \mathbf{P}(H_1) = 0.5.$$

Under H_0, we expect a busy signal with certainty:

$$\mathbf{P}(B \mid H_0) = 1.$$

Under H_1, the conditional probability of B is

$$\mathbf{P}(B \mid H_1) = 0.01.$$

Example IV

Problem : Artemisia moves to a new house and she is "fifty-percent sure" that the phone number is 2537267. To verify this. she uses the house phone to dial 2537267. she obtains a busy signal. and concludes that this is indeed the correct number. Assuming that the probability of a typical seven-digit phone number being busy at any given time is 1%. what is the probability that Artemisia's conclusion was correct?

H_0 : the phone number is 2537267,

H_1 : the phone number is not 2537267,

$$\mathbf{P}(H_0) = \mathbf{P}(H_1) = 0.5.$$

Under H_0, we expect a busy signal with certainty:

$$\mathbf{P}(B \mid H_0) = 1.$$

Under H_1, the conditional probability of B is

$$\mathbf{P}(B \mid H_1) = 0.01.$$

$$\mathbf{P}(H_0 \mid B) = \frac{\mathbf{P}(B \mid H_0)\mathbf{P}(H_0)}{\mathbf{P}(B \mid H_0)\mathbf{P}(H_0) + \mathbf{P}(B \mid H_1)\mathbf{P}(H_1)} = \frac{0.5}{0.5 + 0.005} \approx 0.99.$$

Multiple Hypothesis Testing

- Arises when multiple results are produced and multiple statistical tests are performed
- The tests studied so far are for assessing the evidence for the null (and perhaps alternative) hypothesis for a single result
- A regular statistical test does not suffice
  - For example, getting 10 heads in a row for a fair coin is unlikely for one such experiment
    - probability $= \left(\frac{1}{2}\right)^{10} = 0.001$
  - But, for 10,000 such experiments we would expect 10 such occurrences

Summarizing the Results of Multiple Tests

- The following confusion table defines how results of multiple tests are summarized
  - We assume the results fall into two classes, **+** and **–**, which, follow the alternative and null hypotheses, respectively.
  - The focus is typically on the number of false positives (FP), i.e., the results that belong to the null distribution (– class) but are declared significant (+ class).

Confusion table for summarizing multiple hypothesis testing results.

	Declared significant (+ prediction)	**Declared not significant (– prediction)**	**Total**
H_1 **True (actual +)**	True Positive (TP)	False Negative (FN) type II error	Positives (m_1)
H_0 **True (actual –)**	False Positive (FP) type I error	True Negative (TN)	Negatives (m_0)
	Positive Predictions (Ppred)	Negative Predictions (Npred)	m

Family-wise Error Rate

- By family, we mean a collection of related tests
- **family-wise error rate** (FWER) is the probability of observing even a single false positive (type I error) in an entire set of *m* results.
  - FWER = P(FP > 0).
- Suppose your significance level is 0.05 for a single test
  - Probability of no error for one test is $1 - 0.05 = 0.95$.
  - Probability of no error for *m* tests is 0.95^m
  - FWER = P(FP > 0) = 1 – 0.95^m
  - If *m* = 10, FWER = 0.60

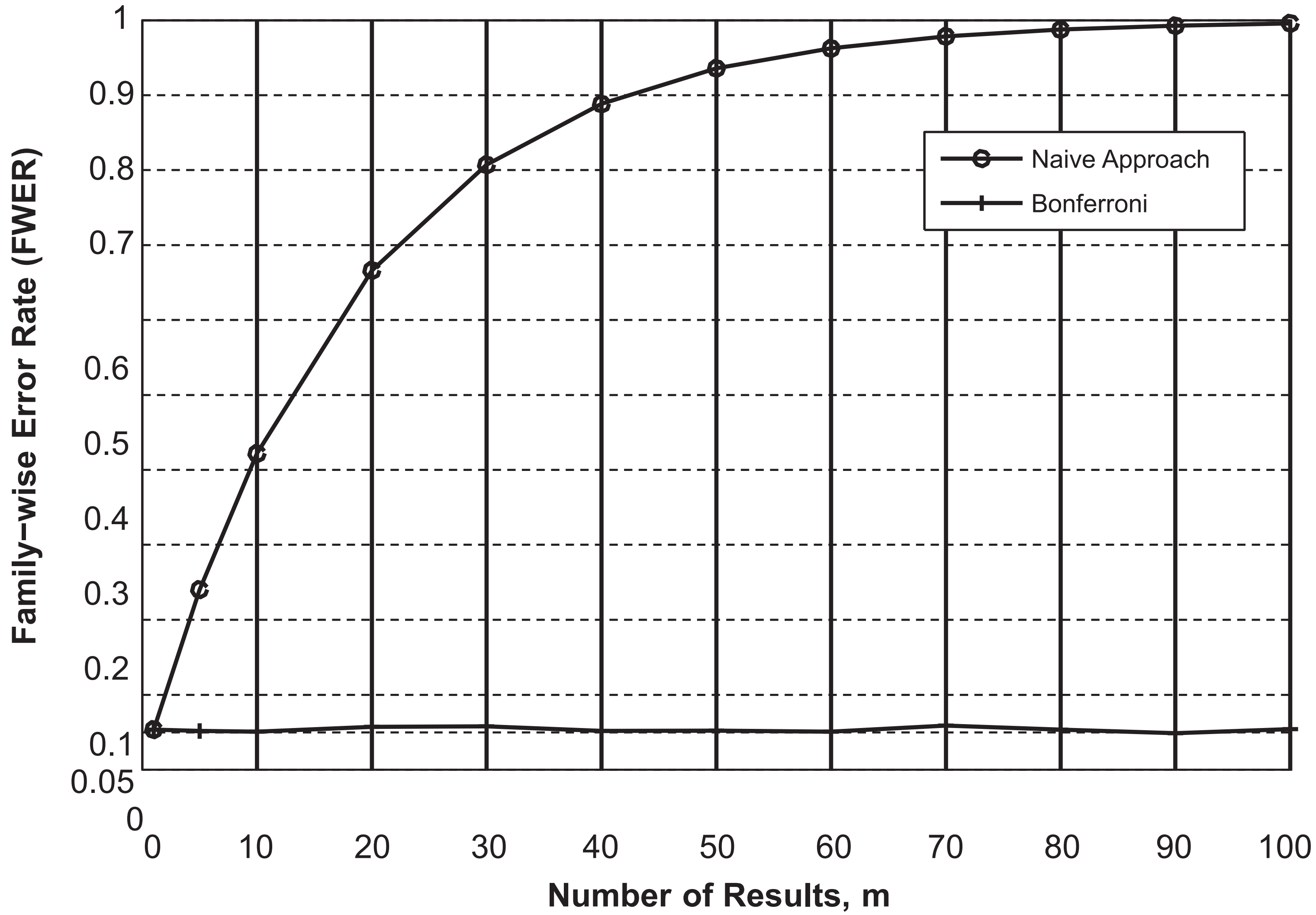
Bonferroni Procedure

- Goal of FWER is to ensure that FWER < α, where α is often 0.05
- Bonferroni Procedure:
  - m results are to be tested
  - Require FWER < α
  - set the significance level, α^* for every test to be $\alpha^* = \alpha/m$.

- If m = 10 and $\alpha = 0.05$ then $\alpha^* = 0.05/10{=}0.005$

Example: Bonferroni versus Naïve approach

- Naïve approach is to evaluate statistical significance for each result without adjusting the significance level.



The family wise error rate (FWER) curves for the naïve approach and the Bonferroni procedure as a function of the number of results, m. $\alpha = 0.05$.

False Discovery Rate

- FWER controlling procedures seek a low probability for obtaining any false positives
  - Not the appropriate tool when the goal is to allow some false positives in order to get more true positives
- The **false discovery rate (FDR)** measures the rate of false positives, which are also called false discoveries

$$Q = \frac{FP}{Ppred} = \frac{FP}{TP + FP} \text{ if } Ppred > 0$$
$$= 0 \text{ if } Ppred = 0,$$

where $Ppred$ is the number of predicted positives

- Thus, FDR = Q P($Ppred$>0) = E(Q), the expected value of Q.

Benjamini-Hochberg Procedure

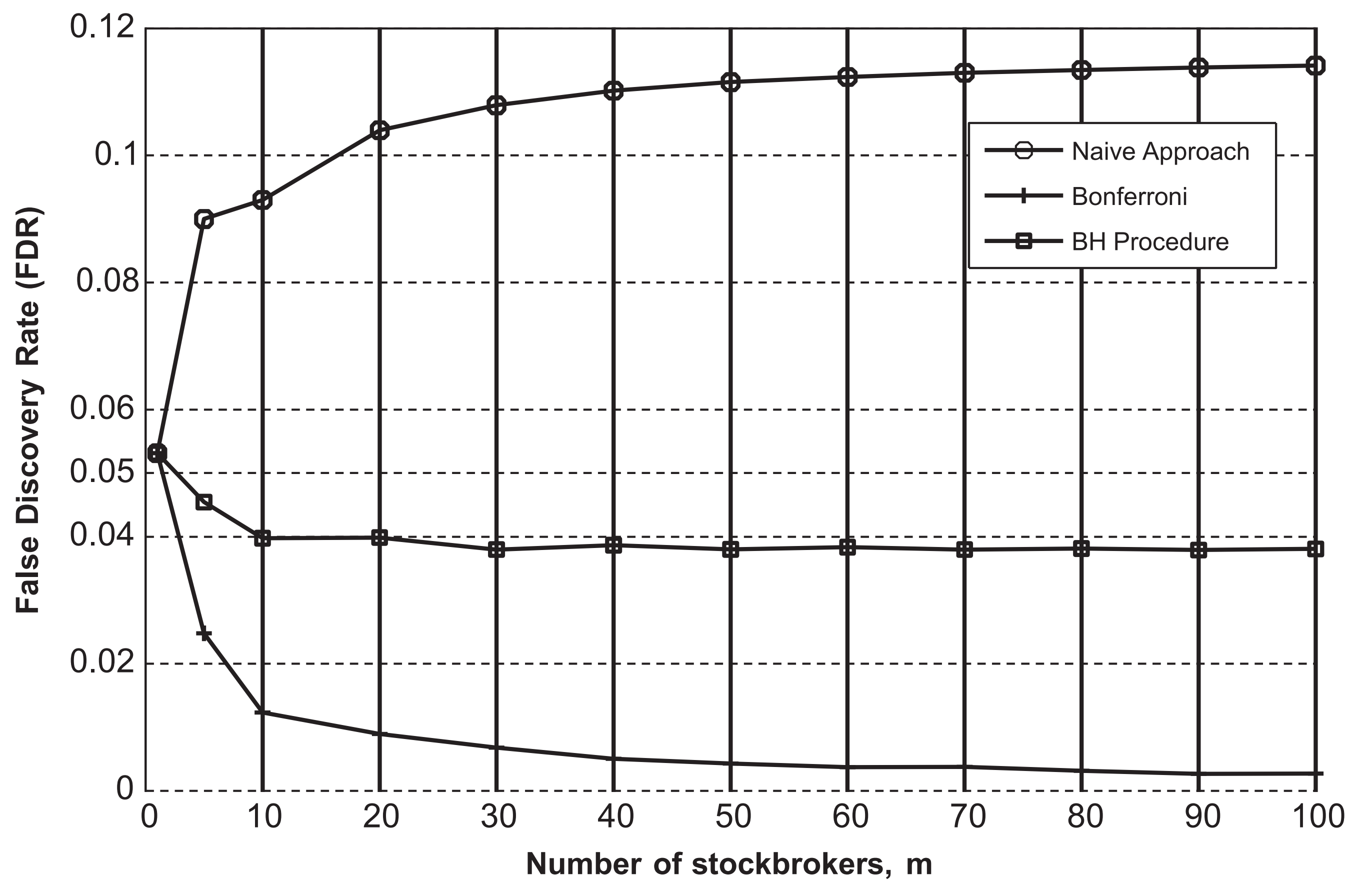
- An algorithm to control the false discovery rate (FDR)

Benjamini-Hochberg (BH) FDR algorithm.
1: Compute p-values for the m results.
2: Order the p-values from smallest to largest (p_1 to p_m).
3: Compute the significance level for p_i as $\alpha_i = i \times \frac{\alpha}{m}$.
4: Let k be the largest index such that $p_k \leq \alpha_k$.
5: Reject H_0 for all results corresponding to the first k p-values, p_i, $1 \leq i \leq k$.

- This procedure first orders the p-values from smallest to largest
- Then it uses a separate significance level for each test
  - $\alpha_i = i \times \frac{\alpha}{m}$

FDR Example: Picking a stockbroker

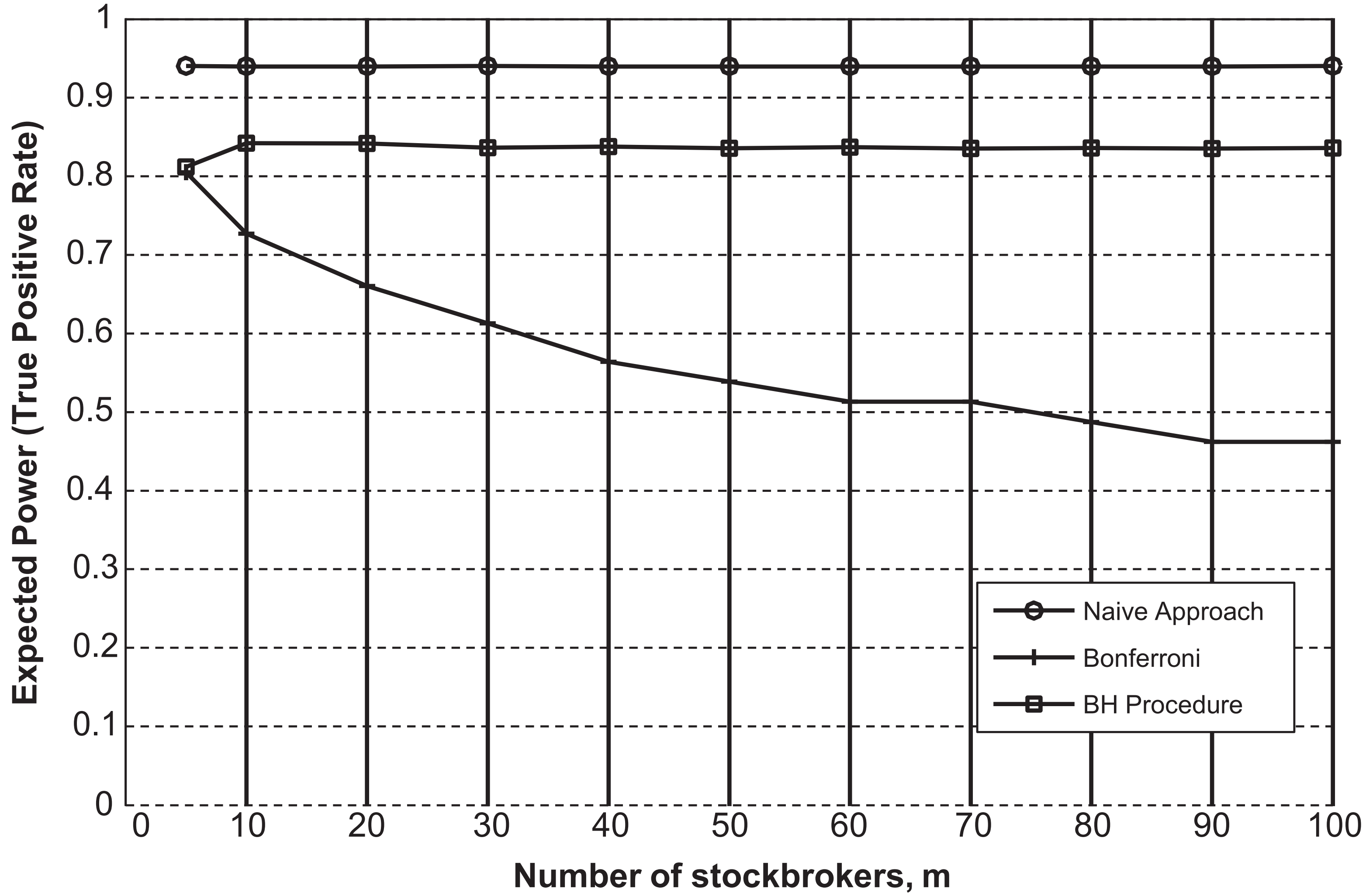
- Suppose we have a test for determining whether a stockbroker makes profitable stock picks. This test, applied to an individual stockbroker, has a significance level, $\alpha = 0.05$. We use the same value for our desired false discovery rate.
  - Normally, we set the desired FDR rate higher, e.g., 10% or 20%
- The following figure compares the naïve approach, Bonferroni, and the BH FDR procedure with respect to the false discovery rate for various numbers of tests, *m*. 1/3 of the sample were from the alternative distribution.



False Discovery Rate as a function of *m*.

FDR Example: Picking a stockbroker …

- The following figure compares the naïve approach, Bonferroni, and the BH FDR procedure with respect to the power for various numbers of tests, *m*. 1/3 of the sample were from the alternative distribution.



Expected Power as function of *m*.

Comparison of FWER and FDR

- FWER is appropriate when it is important to avoid any error.
  - But an FWER procedure such as Bonferroni makes many Type II errors and thus, has poor power.
  - An FWER approach has a very small false discovery rate
- FDR is appropriate when it is important to identity positive results, i.e., those belonging to the alternative distribution.
  - By construction, the false discovery rate is good for an FDR procedure such as the BH approach
  - An FDR approach also has good power

Confidence Intervals

- An estimate $\hat{\Theta}_n$ may not be informative enough
- An $1-\alpha$ **confidence interval** is a (random) interval $[\hat{\Theta}_n^-, \hat{\Theta}_n^+]$,

  s.t. $$\mathbf{P}(\hat{\Theta}_n^- \leq \theta \leq \hat{\Theta}_n^+) \geq 1-\alpha, \quad \forall\ \theta$$

  - often $\alpha = 0.05$, or 0.25, or 0.01
  - interpretation is subtle

Recall: Sample Mean

- $X_1, \ldots, X_n$: i.i.d., mean θ, variance σ^2

$$\hat{\Theta}_n = \text{sample mean} = M_n = \frac{X_1 + \cdots + X_n}{n}$$

Properties:

- $\mathbf{E}[\hat{\Theta}_n] = \theta \quad$ (unbiased)

Sample Variance

- $X_1, \ldots, X_n$: i.i.d., mean θ, variance σ^2

- Estimate of the mean θ:

$$M_n = \frac{X_1 + \ldots + X_n}{n}.$$

- Suppose that we are also interested in an estimator of the variance $v = \sigma^2$.
- Candidate Estimator:

$$\bar{S}_n^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - M_n)^2.$$

- Is this estimator unbiased i.e. is $\mathbf{E}[\bar{S}_n^2] = v$?

Sample Variance

$$\overline{S}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - M_n)^2,$$

$$\mathbf{E}_{(\theta,v)}\left[\overline{S}_n^2\right] = \frac{1}{n}\mathbf{E}_{(\theta,v)}\left[\sum_{i=1}^{n} X_i^2 - 2M_n\sum_{i=1}^{n} X_i + nM_n^2\right]$$

Sample Variance

$$\overline{S}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - M_n)^2,$$

$$\begin{aligned}
\mathbf{E}_{(\theta,v)}\left[\overline{S}_n^2\right] &= \frac{1}{n}\mathbf{E}_{(\theta,v)}\left[\sum_{i=1}^{n} X_i^2 - 2M_n\sum_{i=1}^{n} X_i + nM_n^2\right] \\
&= \mathbf{E}_{(\theta,v)}\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2M_n^2 + M_n^2\right]
\end{aligned}$$

Sample Variance

$$\overline{S}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - M_n)^2,$$

$$\begin{aligned}
\mathbf{E}_{(\theta,v)}\left[\overline{S}_n^2\right] &= \frac{1}{n}\mathbf{E}_{(\theta,v)}\left[\sum_{i=1}^{n} X_i^2 - 2M_n\sum_{i=1}^{n} X_i + nM_n^2\right] \\
&= \mathbf{E}_{(\theta,v)}\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2M_n^2 + M_n^2\right] \\
&= \mathbf{E}_{(\theta,v)}\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - M_n^2\right]
\end{aligned}$$

Sample Variance

$$\overline{S}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - M_n)^2,$$

$$\begin{aligned}
\mathbf{E}_{(\theta,v)}\left[\overline{S}_n^2\right] &= \frac{1}{n}\mathbf{E}_{(\theta,v)}\left[\sum_{i=1}^{n} X_i^2 - 2M_n\sum_{i=1}^{n} X_i + nM_n^2\right] \\
&= \mathbf{E}_{(\theta,v)}\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2M_n^2 + M_n^2\right] \\
&= \mathbf{E}_{(\theta,v)}\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - M_n^2\right] \\
&= \theta^2 + v - \left(\theta^2 + \frac{v}{n}\right)
\end{aligned}$$

Sample Variance

$$\overline{S}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - M_n)^2,$$

$$\begin{aligned}
\mathbf{E}_{(\theta,v)}\left[\overline{S}_n^2\right] &= \frac{1}{n}\mathbf{E}_{(\theta,v)}\left[\sum_{i=1}^{n} X_i^2 - 2M_n\sum_{i=1}^{n} X_i + nM_n^2\right] \\
&= \mathbf{E}_{(\theta,v)}\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - 2M_n^2 + M_n^2\right] \\
&= \mathbf{E}_{(\theta,v)}\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - M_n^2\right] \\
&= \theta^2 + v - \left(\theta^2 + \frac{v}{n}\right) \\
&= \frac{n-1}{n}v.
\end{aligned}$$

Uh Oh!

Sample Variance: Unbiased Estimator

$$\hat{S}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - M_n)^2 = \frac{n}{n-1}\overline{S}_n^2.$$

$$\mathbf{E}_{(\theta,v)}\left[\hat{S}_n^2\right] =$$

Sample Variance: Unbiased Estimator

$$\hat{S}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - M_n)^2 = \frac{n}{n-1}\overline{S}_n^2.$$

$$\mathbf{E}_{(\theta,v)}\left[\hat{S}_n^2\right] = v$$

An unbiased Estimator!

CI Example: Sample Mean

- Estimate of the mean θ:

$$\hat{\Theta}_n = \frac{X_1 + \ldots + X_n}{n}.$$

$$\mathbf{E}[\hat{\Theta}_n] = \theta, \mathrm{Var}[\hat{\Theta}_n] = \frac{\sigma^2}{n}.$$

- Suppose $X_1, \ldots, X_n$ are i.i.d. $N(\theta, \sigma^2)$
- Then $\hat{\Theta}_n$ is a Normal Random Variable (sum of iid Normal random variables is …)

$$\hat{\Theta}_n \sim N\left(\theta, \frac{\sigma^2}{n}\right)$$

CI Example: Sample Mean

- Estimate of the mean θ:

$$\hat{\Theta}_n = \frac{X_1 + \ldots + X_n}{n}.$$

- If $X_1, \ldots, X_n$ are iid $N(\theta, \sigma^2)$, $\hat{\Theta}_n$ is a Normal Random Variable

$$\hat{\Theta}_n \sim N\left(\theta, \frac{\sigma^2}{n}\right)$$

- Noting that if $X \sim N(\mu, \sigma^2)$, then $\frac{X-\mu}{\sigma} \sim N(0,1)$,

$$\frac{\hat{\Theta}_n - \theta}{\sigma/\sqrt{n}} \sim N(0,1)$$

CI Example: Sample Mean

- CI in estimation of the mean
  $\hat{\Theta}_n = (X_1 + \cdots + X_n)/n$

$$\mathbf{P}\left(\frac{|\hat{\Theta}_n - \theta|}{\sigma/\sqrt{n}} \leq 1.96\right) \approx 0.95$$

$$\mathbf{P}\left(\hat{\Theta}_n - \frac{1.96\,\sigma}{\sqrt{n}} \leq \theta \leq \hat{\Theta}_n + \frac{1.96\,\sigma}{\sqrt{n}}\right) \approx 0.95$$

More generally: let z be s.t.

$$\mathbf{P}(Z \leq z) = 1 - \alpha/2$$

$$\mathbf{P}(|Z| \leq z) = 1 - \alpha$$

$$\mathbf{P}\left(\hat{\Theta}_n - \frac{z\sigma}{\sqrt{n}} \leq \theta \leq \hat{\Theta}_n + \frac{z\sigma}{\sqrt{n}}\right) \approx 1 - \alpha$$

The case of unknown variance

- Option 1: use upper bound on σ
  - if X_i Bernoulli: $\sigma \leq 1/2$
- Option 2: use ad hoc estimate of σ
  - if X_i Bernoulli(θ): $\hat{\sigma} = \sqrt{\hat{\Theta}(1 - \hat{\Theta})}$
- Option 3: Use generic estimate of the variance

$$\widehat{S}_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \hat{\Theta}_n)^2$$

(unbiased: $\mathbf{E}[\widehat{S}_n^2] = \sigma^2$)

Example I: Polling

Suppose that we wish to estimate the fraction θ of voters who support a particular candidate for office. We collect n independent sample voter responses $X_1, \ldots, X_n$, where X_i is viewed as a Bernoulli random variable, with $X_i = 1$ if the ith voter supports the candidate. We estimate θ with the sample mean $\hat{\Theta}_n$, and construct a confidence interval based on a normal approximation and different ways of estimating or approximating the unknown variance. For concreteness, suppose that 684 out of a sample of $n = 1200$ voters support the candidate, so that $\hat{\Theta}_n = 684/1200 = 0.57$.

Example I

With the variance estimate:

$$\begin{aligned}\hat{S}_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{\Theta}_n)^2 \\ &= \frac{1}{1199}\left(684\cdot\left(1-\frac{684}{1200}\right)^2 + (1200-684)\cdot\left(0-\frac{684}{1200}\right)^2\right) \\ &\approx 0.245,\end{aligned}$$

we obtain the 95% confidence interval

$$\begin{aligned}\left[\hat{\Theta}_n - 1.96\,\frac{\hat{S}_n}{\sqrt{n}},\ \hat{\Theta}_n + 1.96\,\frac{\hat{S}_n}{\sqrt{n}}\right] &= \left[0.57 - \frac{1.96\cdot\sqrt{0.245}}{\sqrt{1200}},\ 0.57 + \frac{1.96\cdot\sqrt{0.245}}{\sqrt{1200}}\right] \\ &= [0.542,\ 0.598].\end{aligned}$$

Example I

With the variance estimate:

$$\hat{\Theta}_n(1-\hat{\Theta}_n) = \frac{684}{1200} \cdot \left(1 - \frac{684}{1200}\right) = 0.245$$

95% confidence interval

$$\left[\hat{\Theta}_n - 1.96\,\frac{\sqrt{\hat{\Theta}_n(1-\hat{\Theta}_n)}}{\sqrt{n}},\ \hat{\Theta}_n + 1.96\,\frac{\sqrt{\hat{\Theta}_n(1-\hat{\Theta}_n)}}{\sqrt{n}}\right]$$

is again $[0.542,\ 0.598]$.

Example I

With the variance estimate as the upper bound of 1/2:

$$\left[\hat{\Theta}_n - 1.96\,\frac{1/2}{\sqrt{n}},\ \hat{\Theta}_n + 1.96\,\frac{1/2}{\sqrt{n}}\right] = \left[0.57 - \frac{1.96 \cdot (1/2)}{\sqrt{1200}},\ 0.57 + \frac{1.96 \cdot (1/2)}{\sqrt{1200}}\right]$$
$$= [0.542,\ 0.599],$$

which is only slightly wider, but practically the same as before.

Example II

- X: exponential with parameter θ
- Analyze $[a/X, b/X]$
  as a confidence interval for θ:

X is said to have exponential distribution with param. θ if

$$f_X(x;\theta) = \begin{cases} \theta e^{-\theta x} & x \geq 0 \\ 0 & x < 0. \end{cases}$$

Example II

- X: exponential with parameter θ
- Analyze $[a/X, b/X]$
  as a confidence interval for θ:

X is said to have exponential distribution with param. θ if

$$f_X(x;\theta) = \begin{cases} \theta e^{-\theta x} & x \geq 0 \\ 0 & x < 0. \end{cases}$$

$$\begin{aligned} \mathbf{P}\left(\frac{a}{X} \leq \theta \leq \frac{b}{X}\right) &= \mathbf{P}\left(\frac{a}{\theta} \leq X \leq \frac{b}{\theta}\right) \\ &= \int_{a/\theta}^{b/\theta} \theta e^{-\theta x}\, dx \\ &= -e^{-\theta x}\Big|_{x=a/\theta}^{x=b/\theta} = e^{-a} - e^{-b} \end{aligned}$$

Example II

- X: exponential with parameter θ

- Analyze $[a/X, b/X]$
  as a confidence interval for θ:

$$\mathbf{P}\left(\frac{a}{X} \le \theta \le \frac{b}{X}\right) = \mathbf{P}\left(\frac{a}{\theta} \le X \le \frac{b}{\theta}\right)$$
$$= \int_{a/\theta}^{b/\theta} \theta e^{-\theta x}\, dx$$
$$= -e^{-\theta x}\Big|_{x=a/\theta}^{x=b/\theta} = e^{-a} - e^{-b}$$

  No dependence on θ,
  so have a confidence interval

  - Example: $\left[\frac{1}{4X}, \frac{4}{X}\right]$ is a 0.76 confidence interval ("76% CI")

X is said to have exponential distribution with param. θ if

$$f_X(x;\theta) = \begin{cases} \theta e^{-\theta x} & x \ge 0 \\ 0 & x < 0. \end{cases}$$