ESE 402/542 Recitation 6: Hypothesis Testing

Topics

- Hypothesis testing overview
- Neyman-Pearson paradigm

Hypothesis Testing Setup

- H_0: null hypothesis (default model)
- H_1: alternative model (default model not true)

Hypothesis Testing Setup

- H_0: null hypothesis (default model)
- H_1: alternative model (default model not true)
- Accept or reject H_0 based on seen data $X_1, \ldots, X_n$

Hypothesis Testing Setup

- H_0: null hypothesis (default model)
- H_1: alternative model (default model not true)
- Accept or reject H_0 based on seen data $X_1, \ldots, X_n$
- Partition the space of all possible observations $\{X_1, \ldots, X_n\}$ into R and $R^\complement$, based on a test statistic

Hypothesis Testing Setup

- H_0: null hypothesis (default model)
- H_1: alternative model (default model not true)
- Accept or reject H_0 based on seen data $X_1, \ldots, X_n$
- Partition the space of all possible observations $\{X_1, \ldots, X_n\}$ into R and $R^\complement$, based on a test statistic
- R contains all the observations where H_0 is rejected, hence R is the rejection region
- Likewise, $R^\complement$ is the acceptance region

Type I/II errors

- For any rejection region R, there are two errors that can happen
- Type I error: reject H_0 when H_0 is true (samples land in R, but H_0 was true)
- Probability of type I error: $\alpha(R) = P(X_1, \ldots, X_n \in R; H_0)$
- Type II error: accept H_0 when H_0 is false (samples land in $R^{\complement}$, but H_0 was false)
- Probability of type II error: $\beta(R) = P(X_1, \ldots, X_n \notin R; H_1)$

Type I/II errors

- Note that in the traditional hypothesis testing set-up, the Type I error $\alpha(R)$ is precisely the experimenter-designated *significance level*. It quantifies the acceptable level of risk in the experiment: "$\alpha = 0.05$ means there is a 5% probability that the null model occurs by chance, which is unlikely enough for me".
- Type II error is not always accessible to the experimenter. Requires knowledge of the actual alternative distribution. This is not known to us when we define an alternative hypothesis like $H_1 = \theta \neq \theta_0$.
- Type II error is often seen as the more "dangerous" error. This is often due to the convention of notation, e.g. $H_0 =$ Patient does not have disease, $H_1 =$ Patient *does* have disease. Accepting the null hypothesis when the patient *does* have the disease is very bad, from a clinical point of view.

Type I/II errors

- However, mathematically, Type I and II errors are not any more or less important than each other. In fact, sometimes they are philosophically equivalent, e.g.

$$H_0 : \mu = \mu_0$$
$$H_1 : \mu = \mu_1,$$

where μ_0 and μ_1 are possible means of a normal distribution. Then $\alpha = \mathbb{P}[H_1|H_0]$, $\beta = \mathbb{P}[H_0|H_1]$. Note that those errors exactly flip if we had set

$$H_0' : \mu = \mu_1$$
$$H_1' : \mu = \mu_0.$$

$\alpha' = \mathbb{P}[H_1'|H_0'] = \mathbb{P}[H_0|H_1] = \beta,$
$\beta' = \mathbb{P}[H_0'|H_1'] = \mathbb{P}[H_1|H_0] = \alpha.$

Significance Testing

There are not always two well-defined alternative hypotheses. Instead, you may have a question like “Is my coin fair?”.

Significance Testing

There are not always two well-defined alternative hypotheses. Instead, you may have a question like "Is my coin fair?".

- Have null hypothesis of the form $H_0 : \theta = \theta_0$. Alternative hypothesis is simply $H_1 : \theta \neq \theta_0$

Significance Testing

There are not always two well-defined alternative hypotheses. Instead, you may have a question like "Is my coin fair?".

- Have null hypothesis of the form $H_0 : \theta = \theta_0$. Alternative hypothesis is simply $H_1 : \theta \neq \theta_0$
- Choose a test statistic $T(X_1, \ldots, X_n)$

Significance Testing

There are not always two well-defined alternative hypotheses. Instead, you may have a question like "Is my coin fair?".

- Have null hypothesis of the form $H_0 : \theta = \theta_0$. Alternative hypothesis is simply $H_1 : \theta \neq \theta_0$
- Choose a test statistic $T(X_1, \ldots, X_n)$
- Determine the shape of rejection region (where does H_0 get rejected)
- Choose a significance level α_0 where you want the type I error α to be less than

Significance Testing

There are not always two well-defined alternative hypotheses. Instead, you may have a question like "Is my coin fair?".

- Have null hypothesis of the form $H_0 : \theta = \theta_0$. Alternative hypothesis is simply $H_1 : \theta \neq \theta_0$
- Choose a test statistic $T(X_1, \ldots, X_n)$
- Determine the shape of rejection region (where does H_0 get rejected)
- Choose a significance level α_0 where you want the type I error α to be less than
- Determine a threshold t for your statistic T such that the $\alpha \leq \alpha_0$ (this determines rejection region)

Example: Mean of Normal Samples

Suppose we have $X_1, \ldots, X_n$ i.i.d. normal, with mean θ (unknown) and known variance σ^2. We want to know if they are 0-mean or not. Hence $H_0 : \theta = 0$, $H_1 : \theta \neq 0$.

Example: Mean of Normal Samples

Suppose we have $X_1, \ldots, X_n$ i.i.d. normal, with mean θ (unknown) and known variance σ^2. We want to know if they are 0-mean or not. Hence $H_0 : \theta = 0$, $H_1 : \theta \neq 0$.

- Choose $T = \frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}}$

Example: Mean of Normal Samples

Suppose we have $X_1, \ldots, X_n$ i.i.d. normal, with mean θ (unknown) and known variance σ^2. We want to know if they are 0-mean or not. Hence $H_0 : \theta = 0$, $H_1 : \theta \neq 0$.

- Choose $T = \frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}}$
- Naturally, rejection region might be of the form $|T| \geq t$

Example: Mean of Normal Samples

Suppose we have $X_1, \ldots, X_n$ i.i.d. normal, with mean θ (unknown) and known variance σ^2. We want to know if they are 0-mean or not. Hence $H_0 : \theta = 0$, $H_1 : \theta \neq 0$.

- Choose $T = \frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}}$
- Naturally, rejection region might be of the form $|T| \geq t$
- We want to find t such that $\alpha_0 = P(|T| \geq t; \theta = 0)$

Example: Mean of Normal Samples

Suppose we have $X_1, \ldots, X_n$ i.i.d. normal, with mean θ (unknown) and known variance σ^2. We want to know if they are 0-mean or not. Hence $H_0 : \theta = 0$, $H_1 : \theta \neq 0$.

- Choose $T = \frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}}$
- Naturally, rejection region might be of the form $|T| \geq t$
- We want to find t such that $\alpha_0 = P(|T| \geq t; \theta = 0)$
- $P(|T| \geq t; \theta = 0) = P(|\mathcal{N}(0, 1)| \geq t) = \alpha$

Example: Mean of Normal Samples

Suppose we have $X_1, \ldots, X_n$ i.i.d. normal, with mean θ (unknown) and known variance σ^2. We want to know if they are 0-mean or not. Hence $H_0 : \theta = 0$, $H_1 : \theta \neq 0$.

- Choose $T = \frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}}$
- Naturally, rejection region might be of the form $|T| \geq t$
- We want to find t such that $\alpha_0 = P(|T| \geq t; \theta = 0)$
- $P(|T| \geq t; \theta = 0) = P(|\mathcal{N}(0,1)| \geq t) = \alpha$
- Thus, we set $t = z(\alpha/2)$, and reject H_0 if $|T| > t$

Example Problem: Means of Two Populations

Let $X_1, \ldots, X_{n_1}$ be i.i.d. Bernoulli(θ_X) and $Y_1, \ldots, Y_{n_2}$ be i.i.d. Bernoulli(θ_Y). Consider the two hypotheses $H_0 : \theta_X = \theta_Y$ and $H_1 : \theta_X \neq \theta_Y$. Determine a rejection region for H_0 at the α significance level.
Hint 1: Under H_0, $\theta_X = \theta_Y$.
Hint 2: Use a statistic that is based on $\bar{\theta}_X - \bar{\theta}_Y$ and apply the CLT.

Example Problem: Means of Two Populations

Let $X_1, \ldots, X_{n_1}$ be i.i.d. Bernoulli(θ_X) and $Y_1, \ldots, Y_{n_2}$ be i.i.d. Bernoulli(θ_Y). Consider the two hypotheses $H_0 : \theta_X = \theta_Y$ and $H_1 : \theta_X \neq \theta_Y$. Determine a rejection region for H_0 at the α significance level.
Hint 1: Under H_0, $\theta_X = \theta_Y$.
Hint 2: Use a statistic that is based on $\bar{\theta}_X - \bar{\theta}_Y$ and apply the CLT.
Solution: We want to reject H_0 if $\bar{\theta}_X - \bar{\theta}_Y$ is large. Under H_0,

$$\hat{\sigma}^2 = \mathrm{var}\,(\bar{\theta}_X - \bar{\theta}_Y) = \mathrm{var}\,(\bar{\theta}_X) + \mathrm{var}\,(\bar{\theta}_Y) = \frac{1}{n_1}\bar{\theta}_X(1-\bar{\theta}_X) + \frac{1}{n_2}\bar{\theta}_Y(1-\bar{\theta}_Y)$$

$$E[\bar{\theta}_X - \bar{\theta}_Y] = 0$$

Hence $\frac{1}{\hat{\sigma}}(\bar{\theta}_X - \bar{\theta}_Y) \sim \mathcal{N}(0,1)$ Therefore we reject H_0 if $\frac{1}{\hat{\sigma}}(\bar{\theta}_X - \bar{\theta}_Y) > z(\alpha/2)$

Example Problem: