MH1820 Week 11

1. Confidence Interval for Variance
2. Purpose and Rationale of Hypothesis Tests
3. Examples of Hypothesis Testing

Confidence Interval for Variance

Suppose X_1, ..., X_n i.i.d $\sim N(\mu, \sigma^2)$. Notice that

$$\begin{aligned} S^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 \\ \frac{(n-1)S^2}{\sigma^2} &= \sum_{i=1}^{n}\left(\frac{X_i - \overline{X}}{\sigma}\right)^2 \end{aligned}$$

- Recall that $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ (chi-square distribution with degree of freedom $n-1$) (Week 9).
- We can use this to construct confidence intervals.

$$\mathbb{P}\left(\chi^2_{1-\alpha/2}(n-1) < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{\alpha/2}(n-1)\right) = 1-\alpha.$$

Here, $\mathbb{P}(X > \chi^2_\alpha(r)) = \alpha$, i.e. $\chi^2_\alpha(r)$ is the upper $100(1-\alpha)\%$ point.

Rearranging, the $100(1-\alpha)\%$ confidence interval for σ^2 is

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}$$

Theorem 1 (Confidence Interval for σ^2 of normal distribution)

Let $X_1, \ldots, X_n$ i.i.d $\sim N(\mu, \sigma^2)$. The $100(1-\alpha)\%$ confidence interval for σ^2 is given by

$$\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)},$$

where S^2 is the sample variance.

Example 2

The following are the weights, in decagrams, of 10 packages of grass seed distributed by a certain company:

$$46.4, 46.1, 45.8, 47.0, 46.1, 45.9, 45.8, 46.9, 45.2, 46.0.$$

Find a 95% confidence interval for the variance of the weights of all such packages of grass seed distributed by this company, assuming a normal distribution.

Solution.

$$s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right) = \frac{1}{9}\left(\sum_{i=1}^{10} x_i^2 - 10 \cdot 46.12^2\right) \approx 0.286$$

For 95% confidence interval, we have $\alpha = 0.05$. From the χ^2-table, with degree of freedom $n-1=9$, we have $\chi^2_{0.025}(9) = 19.02$, $\chi^2_{0.975}(9) = 2.700$.

Therefore, a 95% confidence interval for σ^2 is

$$0.135 = \frac{(10-1)(0.286)}{19.02} < \sigma^2 < \frac{(10-1)(0.286)}{2.7} = 0.953.$$

$\square$

Purpose and Rationale of Hypothesis Tests

Hypothesis Testing.

A **statistical hypothesis** is an assertion or conjecture concerning one or more populations. For example:

- An engineer claims that the fraction of defective in a process is 0.10.
- A manufacturer claims that the average saturated fat content in a certain rice cereal does not exceed 1.5 grams per serving.
- A project manager claims that the abrasive wear of Material A exceeds that of Material B by 2. more units

The aim of hypothesis tests is to **decide**, based on the given observations, whether to accept or reject the claim.

Before we begin, consider a criminal trial by jury:

- A jury must decide between two hypotheses.
    - The **null hypothesis** H_0: The defendant is innocent.
    - The **alternative hypothesis** H_1: The defendant is guilty.
- The jury does not know which hypothesis is true. They must make a decision on the basis of the evidence presented.

There are two possible decisions.

- Convicting the defendant is called **rejecting the null hypothesis** in favor of the alternative hypothesis. That is, the jury is saying that there is enough evidence to conclude that the defendant is guilty (the alternative hypothesis).

- If the jury acquits it is stating that there is **not enough evidence to support the alternative hypothesis**. Notice that the jury is not saying that the defendant is innocent, only that there is not enough evidence to support the alternative hypothesis.

Choosing H_0:

Given observations $x_1, \ldots, x_n$, the purpose of a hypothesis test is to determine whether a certain "**interesting effect**" exists.

- H_0 should specify a distribution that is reasonable as a population distribution for the observations under the assumption that **no effect** exists.
- Rejecting H_0 means that the observations provide significant evidence for the effect.
- Not rejecting H_0 means that the observations do not contain significant evidence for the effect.

Procedue for Hypothesis Testing:

- Given are observations $x_1, \ldots, x_n$.
- Formulate **null hypothesis** H_0 describing the population distribution from which observations were drawn.
- Choose significance level α (often $\alpha = 0.05$)
- Choose test statistic $T(X_1, \ldots, X_n)$ that contains information on the parameters involved in H_0 and whose distribution is known under H_0.
- Assuming H_0, compute probability (p-value) to observe $t = T(x_1, \ldots, x_n)$ or something "**at least as extreme as t" (in the direction of rejection of H_0)**.
- If the p-value is smaller than α, reject null hypothesis.

Meaning of “at least as extreme”

- Suppose the statistic $T(X_1, \ldots, X_n)$ is used to test H_0.
- Let $t = T(x_1, \ldots, x_n)$ be the obbserved value of T.
- Let $\mathbb{E}[T]$ be the expectation of T under the assumption that H_0 is true. Often deviation from $\mathbb{E}[T]$ is viewed as evidence against H_0.
- **“at least as extreme as t” (in the direction of rejection of H_0)** means
  - $T \geq t$ (one-sided test)
  - $T \leq t$ (one-sided test)
  - $|T - \mathbb{E}[T]| \geq |t - \mathbb{E}[T]|$ (two-sided test)
- The direction of rejection is determined by the alternative hypothesis H_1.

p-value

- p-value is the probability to observe t or something "at least as extreme as t" assuming H_0 is true.
- If p-value is small, it means that chances of observing what we have observed (assuming H_0 is true) is small.

$$\Longrightarrow \text{ the smaller the } p\text{-value, the less we should believe in } H_0.$$

- The significance level α is the minimum value of this probability that we are willing to accept before perfoming the test.

$$\Longrightarrow \begin{cases} \text{Reject } H_0 & \text{if } p\text{-value} < \alpha \\ \text{Do not reject } H_0 & \text{otherwise.} \end{cases}$$

Examples of Hypothesis Testing

Example 3

A random sample of 100 recorded deaths in the US during the past year showed that an average life span of 71.8 years. Assuming a population standard deviation of 8.9 years, does this seem to indicate that the mean life span today is more than 70 years?

Perform a test with $\alpha = 0.05$ as the significance level.

Solution.

Note: population σ^2 is given. By Central Limit Theorem, the sample mean $\overline{X}$, with $n = 100$, is approximately normal. In particular, the statistic

$$T = \frac{X - \mu}{\sigma/\sqrt{n}} \sim N(0,1).$$

Let μ be the population mean.

- Null Hypothesis H_0: $\mu = 70$ (years)
- Alternative Hypothesis H_1: $\mu > 70$ (years)
- Set $\alpha = 0.05$
- Choose statistic $T = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

The Alternative Hypothesis H_1: $\mu > 70$ suggests that we do a one-sided test with p-value $\mathbb{P}(T \geq t)$.

Compute p-value based on data:

$$t = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}} = \frac{71.8 - 70}{8.9/\sqrt{100}} = 2.02$$

$$\begin{aligned} p\text{-value} &= \mathbb{P}(T \geq t) \\ &= \mathbb{P}(T \geq 2.02) \\ &= 1 - \Phi(2.02) = 1 - 0.9783 = 0.0217 < \alpha. \end{aligned}$$

Decision: Reject H_0 (since the p-value is less than $\alpha = 0.05$). □

Example 4

A manufacturer of sports equipment has developed a new synthetic fishing line that the company claims has a mean breaking strength of 8 kilograms with a standard deviation of 0.5 kilograms.

Test the hypothesis that $\mu = 8$ kilograms against the alternative hypothesis that $\mu \neq 8$ kilograms if a sample of 50 lines is tested and found to have a mean breaking strength of 7.8 kilograms.

Use a 0.01 level of significance.

Solution.

- Null Hypothesis H_0: $\mu = 8$ kg
- Alternative Hypothesis H_1: $\mu \neq 8$ kg
- $\alpha = 0.01$
- Statistic: By CLT

$$T = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

The Alternative Hypothesis H_1: $\mu \neq 8$ suggests that we do a two-sided test with p-value $\mathbb{P}(|T - \mathbb{E}[T]| \geq |t - \mathbb{E}[T]|)$. Notice in our case, $\mathbb{E}[T] = 0$.

Compute p-value based on data:

$$t = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}} = \frac{7.8 - 8}{0.5/\sqrt{50}} = -2.83$$

$$\begin{aligned} p\text{-value} &= \mathbb{P}(|T| \geq |t|) \\ &= \mathbb{P}(|T| \geq 2.83) \\ &= 2(1 - \Phi(2.83)) = 2 \times 0.0023 \approx 0.0046 < \alpha = 0.01. \end{aligned}$$

Decision: Reject H_0 (since the p-value is less than $\alpha = 0.01$). □

Example 5

Suppose that the distribution of X is $Bernoulli(p)$. We shall test the null hypothesis $H_0 : p = 0.5$ against the alternative hypothesis $H_1 : p \neq 0.5$.

Suppose a random sample of $n = 100$ observations yielded $\sum_{i=1}^{100} x_i = 65$. Define a test statistic, calculate the p-value and state your conclusion using a significance level of $\alpha = 0.05$.

Solution.

- $H_0 : p = 0.5$
- $H_1 : p \neq 0.5$
- Define the test statistic:

$$T = \sum_{i=1}^{100} X_i \sim Binomial\left(n, p = 0.5\right).$$

Note that $\mathbb{E}[T] = np = 50$.

Compute p-value based on data:

$$t = \sum_{i=1}^{100} x_i = 65.$$

$$\begin{aligned}
p\text{-value} &= \mathbb{P}(|T - \mathbb{E}[T]| \geq |t - \mathbb{E}[T]|) \\
&= \mathbb{P}(|T - 50| \geq |65 - 50|) \\
&= \mathbb{P}(|T - 50| \geq 15) \\
&= \mathbb{P}(T \geq 65) + \mathbb{P}(T \leq 35) \\
&= 1 - \mathbb{P}(35 < T < 65) \\
&\approx 1 - \left(\Phi\left(\frac{65 - 100(0.5)}{0.5\sqrt{100}}\right) - \Phi\left(\frac{35 - 100(0.5)}{0.5\sqrt{100}}\right)\right)
\end{aligned}$$

by CLT.

$$
\begin{aligned}
p\text{-value} &\approx 1 - \Phi(3) + \Phi(-3) \\
&\approx 1 - 0.9987 + 0.0013 = 0.0026 < \alpha = 0.05.
\end{aligned}
$$

Decision: Reject H_0 (since the p-value is less than $\alpha = 0.05$). $\square$