LECTURE - 23

Agenda:

1. Some large sample tests
2. Examples

We have established the basic definitions for hypothesis testing in the last two lectures. We now consider some standard situations and some common tests that are used in these situations.

Suppose we have data from a population with unknown parameter θ. Let us assume that there exists an estimate $\hat{\theta}$, such that

$$\frac{\hat{\theta} - \theta}{\underbrace{\widehat{SE}(\hat{\theta})}_{\text{Refers to an estimator of } SE(\hat{\theta}).}}$$ has an approximate Normal(0,1) distribution.

Let θ_0 be a fixed number.

TASK 1: Provide a test statistic and rejection region with level α for testing

$$H_0: \theta = \theta_0 \quad \text{v.s.} \quad H_A: \theta > \theta_0$$

ANSWER: TEST STATISTIC: $\hat{\theta}$.

REJECTION REGION: $\hat{\theta} > \theta_0 + z_{1-\alpha} \widehat{SE}(\hat{\theta})$.

(reject for large values of the estimator)

Let us check that the rejection region indeed corresponds to level α.

$$\text{Level} = P(H_0 \text{ is rejected} \mid H_0 \text{ is true})$$

$$= P\left(\hat{\theta} > \theta_0 + z_{1-\alpha} \widehat{SE}(\hat{\theta}) \mid \theta = \theta_0 \right)$$

$$= P\left(\frac{\hat{\theta} - \theta_0}{\widehat{SE}(\hat{\theta})} > z_{1-\alpha} \mid \theta = \theta_0 \right)$$

Note that if $\theta = \theta_0$, then $\frac{\hat{\theta} - \theta_0}{\widehat{SE}(\hat{\theta})}$ has an approximate Normal $(0, 1)$ distribution. Hence,

$$\text{Level} \approx P(\text{Normal}(0,1) > z_{1-\alpha}) = \alpha.$$

$$\boxed{\text{RECALL THAT } z_\alpha \text{ IS THE } \alpha^{th} \text{ QUANTILE OF THE NORMAL}(0,1) \text{ DISTRIBUTION, I.E., FOR EVERY } \alpha, \; P(Z < z_\alpha) = \alpha.}$$

TASK 2: Provide a test statistic and rejection region with level α for testing

$$H_0: \theta = \theta_0 \quad \text{v.s.} \quad H_A: \theta < \theta_0.$$

ANSWER: TEST STATISTIC: $\hat{\theta}$

REJECTION REGION: $\hat{\theta} < \theta_0 + z_\alpha \hat{SE}(\hat{\theta})$.

(reject for small values of the estimator)

Let us check that the rejection region indeed corresponds to level α.

$$\begin{aligned}
\text{Level} &= P(H_0 \text{ is rejected} \mid H_0 \text{ is true}) \\
&= P(\hat{\theta} < \theta_0 + z_\alpha \hat{SE}(\hat{\theta}) \mid \theta = \theta_0) \\
&= P\left(\frac{\hat{\theta} - \theta_0}{\hat{SE}(\hat{\theta})} < z_\alpha \;\middle|\; \theta = \theta_0 \right) \\
&\approx P(\text{Normal}(0,1) < z_\alpha \mid \theta = \theta_0) \\
&= \alpha.
\end{aligned}$$

TASK 3: Provide a test statistic and rejection region with level α for testing

$$H_0: \theta = \theta_0 \quad \text{v.s.} \quad H_A: \theta \neq \theta_0.$$

ANSWER:

TEST STATISTIC: $\hat{\theta}$

REJECTION REGION: $\hat{\theta} < \theta_0 + z_{\alpha/2}\,\widehat{SE}(\hat{\theta})$ or $\hat{\theta} > \theta_0 + z_{1-\alpha/2}\,\widehat{SE}(\hat{\theta})$

(reject for too small or too big values of the estimator)

Let us check that the rejection region indeed corresponds to level α.

$$\text{Level} = P(H_0 \text{ is rejected} \mid H_0 \text{ is true})$$

$$= P\left(\hat{\theta} < \theta_0 + z_{\alpha/2}\,\widehat{SE}(\hat{\theta}) \text{ or } \hat{\theta} > \theta_0 + z_{1-\frac{\alpha}{2}}\,\widehat{SE}(\hat{\theta}) \;\middle|\; \theta = \theta_0\right)$$

$$= P\left(\hat{\theta} < \theta_0 + z_{\alpha/2}\,\widehat{SE}(\hat{\theta}) \mid \theta = \theta_0\right) + P\left(\hat{\theta} > \theta_0 + z_{1-\frac{\alpha}{2}}\,\widehat{SE}(\hat{\theta}) \mid \theta = \theta_0\right)$$

$$= P\left(\frac{\hat{\theta} - \theta_0}{\widehat{SE}(\hat{\theta})} < z_{\alpha/2} \;\middle|\; \theta = \theta_0\right) + P\left(\frac{\hat{\theta} - \theta_0}{\widehat{SE}(\hat{\theta})} > z_{1-\frac{\alpha}{2}} \;\middle|\; \theta = \theta_0\right)$$

$$\approx P\left(\text{Normal}(0,1) < z_{\alpha/2}\right) + P\left(\text{Normal}(0,1) > z_{1-\frac{\alpha}{2}}\right)$$

$$= \frac{\alpha}{2} + \frac{\alpha}{2}$$

$$= \alpha.$$

The alternatives $H_A: \theta > \theta_0$ and $H_A: \theta < \theta_0$ are known as ONE-SIDED ALTERNATIVE HYPOTHESES.
The alternative $H_A: \theta \neq \theta_0$ is known as a TWO-SIDED ALTERNATIVE HYPOTHESIS.

<u>Example:</u> Suppose $Y_1, Y_2, ---, Y_{100}$ are the lifetimes of a random sample of electronic devices from a population of such devices. It is believed that the average lifetime μ of the population is 5 time units. However, an engineer claims that due to recent improvements in production quality, the average lifetime is now more than 5. Suppose that for the observed data, $\bar{Y} = 7$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 = 16$.

Is there evidence to conclude that the engineer is right with level $= 0.05$?

<u>ANSWER:</u> We want to test

$$H_0: \mu = 5 \quad \text{v.s.} \quad H_A: \mu > 5.$$

Let us try and verify whether the conditions required for the abstract situation which we just studied are satisfied.

We first need to look for an estimator $\hat{\mu}$. Clearly, $\hat{\mu} = \bar{Y}$ is the most intuitive and natural estimator.

Note that $SE(\hat{\mu}) = \sqrt{V(\hat{\mu})} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$.

Clearly, an intuitive and natural estimator for σ is S. Hence,

$$\hat{SE}(\hat{\mu}) = \frac{S}{\sqrt{n}}.$$

Also, $\frac{\hat{\mu} - \mu}{\hat{SE}(\hat{\mu})} = \frac{\sqrt{n}(\bar{Y} - \mu)}{S}$ has an approximate Normal(0,1) distribution. Hence,

TEST STATISTIC: $\hat{\mu}$

REJECTION REGION: $\hat{\mu} > 5 + z_{1-0.05}\,\hat{SE}(\hat{\mu})$

$$\Updownarrow$$

$$\hat{\mu} > 5 + \frac{1.68\,S}{\sqrt{100}}$$

$$\Updownarrow$$

$$\hat{\mu} > 5.672.$$

Since the observed $\hat{\mu}$ is $7 > 5.672$, we conclude that there is evidence that the engineer is right.

One can similarly construct tests for $H_A: \mu < 5$ or $H_A: \mu \neq 5$ in this situation.

(Simple home exercise: Work these tests out on your own)