LECTURE - 27

<u>Agenda:</u>

1. p- values
2. Examples

Consider a general hypothesis testing problem concerning two hypotheses H_0 and H_A about an unknown parameter θ. Suppose we decide that $\hat{\theta}$ is the appropriate test statistic. Then, given that we want to test at level α, we construct an appropriate rejection region. Depending on whether $\hat{\theta}_{observed}$ (observed value of $\hat{\theta}$) lies in the rejection region or not, we reject or accept H_0. <u>NOTE THAT THE REJECTION REGION DEPENDS ON THE CHOICE OF THE LEVEL (α).</u>

Hence, it is possible that if Person A and Person B analyze the same data, but Person A chooses level 0.05, whereas Person B chooses level 0.01, they reach different conclusions. Keeping this fact in mind, the statistical community devised a more informative way of reporting the conclusions of a testing procedure. Instead of reporting "Based on the observed data, H_0 was rejected (or accepted) at level-α", we report the "p-value" of the testing procedure.

DEFINITION: If $\hat{\theta}$ is a test statistic, then the p-value of the corresponding test is defined as the <u>smallest level α</u> for which we will reject the null hypothesis based on the observed data.

Hence, if the level of the test is chosen to be larger than the p-value, we would reject the null hypothesis H_0. On the other hand, if the level of the test is chosen to be smaller than the p-value, we would accept the null hypothesis H_0.

THE SMALLER THE p-value, STRONGER IS THE EVIDENCE THAT THE NULL HYPOTHESIS SHOULD BE REJECTED.

Instead of just arbitrarily fixing the level of the test and reporting acceptance or rejection, the p-value allows the client or the reader to evaluate the extent to which the observed data disagree with the null hypothesis.

In some sense, rather than just saying who won or who lost, the p-value also gives us the idea about the margin of victory.

RESULT: If the rejection region is of the form $\hat{\theta} \leq k$, then if $\hat{\theta}_{observed}$ denotes the observed value of $\hat{\theta}$,

$$p\text{-value} = P\left(\hat{\theta} \leq \hat{\theta}_{observed} \mid H_0 \text{ is true}\right)$$

RESULT: If the rejection region is of the form $\hat{\theta} \geq k$, then if $\hat{\theta}_{observed}$ denotes the observed value of $\hat{\theta}$,

$$p\text{-value} = P\left(\hat{\theta} \geq \hat{\theta}_{observed} \mid H_0 \text{ is true}\right)$$

Example: Suppose the mean lifetime μ for a population of bulbs is believed to be 18 time units. A researcher claims that due to improvement in the quality of production, the new mean lifetime is greater than 18. The data $Y_1, Y_2, ---, Y_{100}$ was gathered for the lifetime of 100 randomly chosen bulbs. If the observed $\bar{Y}$ is 20 and S is 14, provide the p-value of the standard testing procedure.

Note that we want to test

$$H_0: \mu = 18 \quad \text{v.s.} \quad H_A: \mu > 18.$$

TEST STATISTIC: $\overline{Y}$

REJECTION REGION: Reject for large values of $\overline{Y}$.

Hence,

$$p\text{-value} = P\left(\overline{Y} > \overline{Y}_{\text{observed}} \,\middle|\, H_0 \text{ is true}\right)$$

$$= P\left(\overline{Y} > 20 \mid \mu = 18\right)$$

$$= P\left(\frac{\overline{Y} - 18}{S} > \frac{20 - 18}{14} \,\middle|\, \mu = 18\right)$$

$$= P\left(\frac{\sqrt{100}\,(\overline{Y} - 18)}{S} > \sqrt{100}\left(\frac{20-18}{14}\right) \,\middle|\, \mu = 18\right)$$

$$\approx P\left(\text{Normal}(0,1) > \frac{10}{7}\right)$$

$$= 1 - \Phi\left(\frac{10}{7}\right).$$