Bruce M. Boghosian

Hypothesis Testing

Testing for inequalty

Testing for equality

Summary

Hypothesis testing and decision rules

Bruce M. Boghosian

School of Arts and Sciences

Department of Mathematics

Tufts University

Tufts UNIVERSITY

Outline

1 Hypothesis Testing

2 Testing for inequalty

3 Testing for equality

4 Summary

The Decision Rule

- Imagine a *yes/no question* about random data.
  - Deciding whether or not a person on trial *is or is not* guilty
  - Testing mileage on cars to decide if a fuel additive *is or is not* effective
- In both examples, there is a *null hypothesis* H_0: This is the thing you would suppose to be true by default – in the absence of evidence to the contrary.
  - In a trial, the null hypothesis in the U.S. is "innocent until proven guilty"
  - In a fuel additive test, the null hypothesis is that the additive is ineffective.
- There is also an *alternative hypothesis* H_1: This might be the thing you are trying to use evidence to prove is true.
  - In a trial, the alternative hypothesis might be that the defendant is guilty.
  - In fuel additive test, alternative hypothesis is that additive is proven effective.

The fuel additive problem

- We have a fleet of n cars whose mileage is $\overline{y} = 25$ mpg.
- Sample mean is *normally distributed* with *known variance* of $\sigma = 2.4$ mpg.
- In test with the additive, we found mileages $y_1, \ldots, y_n$, with *sample mean*

$$\overline{y} = \frac{1}{n} \sum_{j=1}^{n} y_j = 26.3 \text{ mpg}$$

- The competing hypotheses are
  - H_0: $\mu = 25.0$ so the additive is ineffective
  - H_1: $\mu > 25.0$ so the additive is effective
- Note $\overline{y} < 25.0$ is certainly not grounds to reject H_0.
- Even values of $\overline{y}$ slightly greater than 25.0 might not contradict H_0.
- The question is "How large must $\overline{y}$ be before we stop believing H_0?"

The fuel additive problem (continued)

- How large must $\overline{y}$ be before we stop believing H_0?
- There is no right answer to this question.
- It depends how much confidence you want in your answer.
- Suppose we want 95% confidence. Then we seek y^* such that we can reject H_0 whenever $\overline{y} > y^*$, and have 95% confidence that our judgement is correct.

$$
\begin{aligned}
P(\text{We reject } H_0 \mid H_0 \text{ is true}) &= 0.05 \\
\therefore \quad P(\overline{y} > y^* \mid \mu = 25.0) &= 0.05 \\
\therefore \quad P\left(\frac{\mu - \overline{y}}{2.4/\sqrt{30}} < \frac{\mu - y^*}{2.4/\sqrt{30}} \mid \mu = 25.0\right) &= 0.05 \\
\therefore \quad P\left(Z < \frac{\mu - y^*}{2.4/\sqrt{30}}\right) &= 0.05
\end{aligned}
$$

The fuel additive problem (continued)

- We can have the desired confidence if

$$P\left(Z < \frac{\mu - y^*}{2.4/\sqrt{30}}\right) = 0.05$$

$$\therefore \quad \frac{25.0 - y^*}{2.4/\sqrt{30}} = z_{0.95} = -1.64$$

$$\therefore \quad \frac{y^* - 25.0}{2.4/\sqrt{30}} = z_{0.05} = +1.64$$

$$\therefore \quad y^* = 25.0 + 1.64\frac{2.4}{\sqrt{30}} = 25.7186$$

- So we should reject the null hypothesis if $\overline{y} > 25.7186$.
- In this case, since $\overline{y} = 26.3$, we reject the null hypothesis.

Testing for equality

- Suppose we wish to test $H_0 : \ \mu = \mu_0$.
- This might have been $\mu_0 = 25$ in the previous example.
- We want 95% confidence that we are right, so $\alpha = 0.05$.
- Suppose we accept H_0 if $\mu - \Delta y \leq \overline{y} \leq \mu + \Delta y$, and reject otherwise. So demand

$$\begin{aligned}
P(\text{We reject } H_0 \mid H_0 \text{ is true}) =& \alpha \\
\therefore \quad P(\overline{y} < \mu - \Delta y \mid \mu = \mu_0) + P(\overline{y} > \mu + \Delta y \mid \mu = \mu_0) =& \alpha \\
\therefore \quad P(\overline{y} > \mu + \Delta y \mid \mu = \mu_0) =& \alpha/2 \\
\therefore \quad P\left(\frac{\overline{y} - \mu}{\sigma/\sqrt{n}} > \frac{\Delta y}{\sigma/\sqrt{n}} \mid \mu = \mu_0\right) =& \alpha/2
\end{aligned}$$

- We know $Z = \frac{\overline{y}-\mu}{\sigma/\sqrt{n}}$ is distributed as a standard normal.
- Hence $\frac{\Delta y}{\sigma/\sqrt{n}} + z_{\alpha/2}$, or

$$\Delta y = +z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Summary of the above

- Easier to state in terms of $z := \frac{\overline{y}-\mu_0}{\sigma/\sqrt{n}}$
- Let $y_1, \ldots, y_n$ be a random sample from a normal distribution for which σ is known.
- To test $H_0: \ \mu = \mu_0$ versus $H_1: \ \mu > \mu_0$ at the α level of significance, reject H_0 if $z \geq z_\alpha$.
- To test $H_0: \ \mu = \mu_0$ versus $H_1: \ \mu < \mu_0$ at the α level of significance, reject H_0 if $z \leq -z_\alpha$.
- To test $H_0: \ \mu = \mu_0$ versus $H_1: \ \mu \neq \mu_0$ at the α level of significance, reject H_0 if either $z \leq -z_{\alpha/2}$ or $z \geq +z_{\alpha/2}$.

Caveats and pitfalls

- All of the above tests are designed to accept or reject the null hypothesis.
- The above analysis worked because assuming the null hypothesis H_0 means assuming that $\mu = \mu_0$. We could then subtract μ_0 across the inequality and divide by the variance to obtain a random variable distributed like a standard normal.
- We do not have a way to accept or reject the alternative hypothesis.
- If H_0 is not rejected, it is better to characterize that as "failure to reject the null hypothesis," rather than "accept the null hypothesis".

Summary

- We discussed null and alternate hypotheses.
- We discussed hypothesis testing and decision rules
  - for *inequality* involving normally distributed r.v.
  - for *equality* involving normally distributed r.v.