Statistical testing: introduction

Josemari Sarasola

Statistics for Business

Gizapedia

Statistical testing: introduction

Example: Planet Nine

- 2016: Astronomers Mike Brown and Konstantin Batygin inferred the existence of a planet beyond Neptune (and Pluto): Planet Nine.
- Take a null hypothesis as a basis (H_0): there is no planet beyond Neptune (let's be cautious: we have not seen anything).
- Evidence: Brown and Batygin saw that the orbits of some TNOs (Trans Neptunian Obejct) were clustered.
- Assumed H_0, there is a very small probability of that occurring. Alternatively, there is a big probability for that if Planet Nine really exists.
- Conclusion: we reject there is no Planet Nine, so we decide it must be somewhere. But we cannot prove that until we find it!

Steps

- Fix the null hypothesis: H_0 .
- We'll take into account these criteria to fix H_0, in this priority order:
  1. Take the hypothesis that is given.
  2. Take the opposite of that what it's asked or thought.
  3. Take the opposite of that what data are showing us.
- Calculate the probability of that what occurred (or something stranger), that is the p-value, assumed H_0.
- To calculate that probability take the boundary value in H_0.
- Compare p-value to α value fixed in advance.
  - $p \leq \alpha$ implies that what happened is rare, and so we reject H_0.
  - $p > \alpha$ implies that what happened is normal and so we accept H_0.

Statistical testing: introduction

Example: p-value

Normally, the probability of a piece being faulty is 0.15. We have bought a new machine and think that probability has decreased. Among 20 pieces, only one has been faulty. Should we conclude that the machine improves the process? $\alpha = 0.1$.

- $H_0 : p \geq 0.15$ (the opposite of that we are asking)
- We will reject $H_0 : p \geq 0.15$ when there are *few* faulty pieces among the set of 20 pieces (critical region is on the lower side).
- We take the boundary value ($p = 0.15$) as the basis.
- X: faulty pieces among the set of 20 pieces.
  $P[X \leq 1/B(n = 20, p = 0.15)] = 0.1755$.
- p-value (0.1755)is bigger than α. So, we decide that the evidence is normal, assumed H_0, and accordingly we accept H_0.

Example: critical value

Which is the critical value for the previous test? Critical value is the smallest value in order to reject the null hypothesis.

- In our example, $P[X \leq 0/B(n = 20, p = 0.15)] = 0.0387$.
- So, if there had been 0 faulty pieces, we should have rejected the null hypothesis. As there are more faulty pieces (actually, 1), we accept H_0.
- So, there are to ways to test a hypothesis:
  - Comparing p-value to α.
  - Comparing evidence to the critical value.
- With the critical value we can design a test without having any evidence. But it's not possible to calculate the p-value without any data.