Hypothesis testing and decision rules

Testing binomial data

Bruce M. Boghosian

Tufts UNIVERSITY | School of Arts and Sciences

Department of Mathematics

Tufts University

Outline

Single Bernoulli trial

- We have $Y \in \{0, 1\}$ ($0 =$ tails and $1 =$ heads).
- Bernoulli distribution for one trial $p_Y(k; p) = p^k(1-p)^{1-k}$

$$p_Y(0; p) = p^0(1-p)^{1-0} = 1 - p$$
$$p_Y(1; p) = p^1(1-p)^{1-1} = p$$

- We have $E(Y) = p$ and $\mathrm{Var}(Y) = p(1-p)$ for one trial
- Parameter p is unknown.

Multiple Bernoulli trials – the Binomial Distribution

- We conduct n independent Bernoulli trials
- Random vars. $\vec{Y} = \langle Y_1, Y_2, \ldots, Y_n \rangle$ where $Y_j \in \{0, 1\}$.
- Outcomes $\vec{k} = \langle k_1, k_2, \ldots, k_n \rangle$ where $k_j \in \{0, 1\}$.
- Random var. for total heads is $Y := \sum_{j=1}^{n} Y_j = n\overline{Y}$
- Outcome for total number of heads is $k := \sum_{j=1}^{n} k_j = n\overline{k}$
- Binomial distribution for n trials is

$$p_Y(k; p) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{n\overline{k}} p^{n\overline{k}} (1-p)^{n(1-\overline{k})}.$$

- We have $E(Y) = np$ and $\mathrm{Var}(Y) = np(1-p)$ for n trials
- Parameter p is unknown.

Large-sample versus small-sample testing

- Note $0 \leq Y \leq n$ and $E(Y) = np$ and $\sigma_Y = \sqrt{np(1-p)}$
- If n is sufficiently large, $[pn - 3\sigma_Y, pn + 3\sigma_Y] \subset [0, n]$
- Because p is unknown, we use H_0 to make this judgement
- We do a *large-sample test*, relying on CLT, if

$$\begin{aligned} 0 <& np_0 - 3\sqrt{np_0(1-p_0)} < np_0 + 3\sqrt{np_0(1-p_0)} < n. \\ \therefore\ 0 <& p_0 - 3\sqrt{p_0(1-p_0)/n} < p_0 + 3\sqrt{p_0(1-p_0)/n} < 1. \end{aligned}$$

- If above true, normal distribution f_Y obtained from CLT comfortably fits in $[0, n]$ (respectively $f_{\overline{Y}}$ fits in $[0, 1]$)
- If above not true, we must conduct a *small-sample test*.

Large-sample testing

Bruce M. Boghosian

- Suppose Z is distributed as a standard normal, where

$$Z := \frac{Y - np_0}{\sqrt{np_0(1-p_0)}} = \frac{\overline{Y} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

- Take n samples, $k_1, \ldots, k_n$, let $k = \sum_j^n k_j$, and define

$$z := \frac{k - np_0}{\sqrt{np_0(1-p_0)}} = \frac{\overline{k} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

- Use hypothesis testing method for standard normal r.v.
- Find thresholds for not rejecting H_0 if it is true with $100(1-\alpha)\%$ confidence.

Large-sample testing (continued)

- Take n samples, $k_1, \ldots, k_n$, let $k = \sum_j^n k_j$, and define

$$z := \frac{k - np_0}{\sqrt{np_0(1-p_0)}} = \frac{\overline{k} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

- Three tests similar to earlier work on hypothesis testing:
    - To test $H_0 : p = p_0$ versus $H_1 : p > p_0$ at the α level of significance, reject H_0 if $z \geq +z_\alpha$
    - To test $H_0 : p = p_0$ versus $H_1 : p < p_0$ at the α level of significance, reject H_0 if $z \leq -z_\alpha$
    - To test $H_0 : p = p_0$ versus $H_1 : p \neq p_0$ at the α level of significance, reject H_0 if z is either $\leq -z_{\alpha/2}$ or $\geq +z_{\alpha/2}$.

Example: Do people postpone their deaths?

- Study of $n = 747$ obituaries
- Only 60, or 8%, died during the three months prior to their birthday.
- If deaths were distributed uniformly, one would expect this figure to be 25%.
- Is the decrease from 25% to 8% statistically significant?

Example: Death postponement (continued)

- Define k_j, for $j = 1, \ldots, 747$ to be
  - $= 1$ if jth person died in three months prior to birthday
  - $= 0$ if jth person died at any other time of the year
- Then $p_e = \overline{k} = \frac{1}{n}\sum_j^n k_j$ is fraction of deaths three months prior to a birthday
- Take $H_0 : p = 0.25$ since the contrary seems perverse.
- Take $H_1 : p < 0.25$ and demand confidence with $\alpha = 0.05$.

Example: Death postponement (continued)

- First note that $np_0 = 747(0.25) = 186.75$ and

$$\sigma = \sqrt{747(0.25)(1-0.25)} = 11.83.$$

- np_0 more than 3σ greater than zero and less than $n = 747$.
- Hence large-sample testing is warranted.
- Null hypothesis is that $p = p_0 = 0.25$, and $n = 747$
- Calculate

$$z = \frac{k - np_0}{\sqrt{np_0(1-p_0)}} = \frac{60 - 747(0.25)}{\sqrt{747(0.25)(1-0.25)}} = -10.7.$$

- The above is far, far less than $-z_{0.05} = -1.64$.
- Very strong evidence ($\gg 95\%$ confidence) effect is real.

Small-sample testing – explained by example

- Example: Experimental drug test with $n = 19$ patients
- Old treatment is known to be 85% effective
- We wish to test $H_0 : p = 0.85$ versus $H_1 : p \neq 0.85$
- For $n = 19$ and $p_0 = 0.85$,
  - $np_0 = 16.15$
  - $\sigma = \sqrt{19(0.85)(0.15)} = 1.556$
- Note that $np_0 + 3\sigma = 16.15 + 3(1.556) = 20.819 > 19$
- Indicates that small-sample testing is necessary

How does small-sample testing work?

- List 19 possibilities:

k	$P(Y=k)=\binom{19}{k}(0.85)^k(0.15)^{19-k}$
6	1.99151×10^{-7}
7	2.09582×10^{-6}
8	0.0000178145
9	0.000123382
10	0.000699164
11	0.00324158
12	0.012246
13	0.0373659
14	0.0907457
15	0.171409
16	0.242829
17	0.242829
18	0.152892
19	0.0455994

- Note $P(Y \leq 13) = 0.053\ldots$, and $P(Y = 19) = 0.045\ldots$
- Hence we reject H_0 if $k \leq 13$ or $k = 19$.
- Note that confidence interval is asymmetric.

Summary

- We have studied the testing of binomial data
- We have described two cases
  - Large-sample testing
  - Small-sample testing
- We have provided an example of each kind of testing.