Two-sample inferences I

Bruce M. Boghosian

Department of Mathematics

Tufts University

Outline

The one-sample t test is a GLRT

- Recall the one-sample t test wherein we calculate $t = \frac{\overline{y}-\mu_0}{s/\sqrt{n}}$, and draw conclusions such as: To test $H_0 : \mu = \mu_0$ versus $H_1 : \mu > \mu_0$ at the α level of significance, reject H_0 if $t \geq t_{\alpha,n-1}$, etc.
- Recall the definition of a GLRT: Reject H_0 whenever $0 < \lambda \leq \lambda^*$, where $\lambda = \frac{\max_\omega L(\theta_1,\ldots,\theta_k)}{\max_\Omega L(\theta_1,\ldots,\theta_k)}$, and λ^* is chosen so that $P\left(0 < \Lambda \leq \lambda^* \mid H_0 \text{ is true}\right) = \alpha$
- In the present circumstance

$$\omega = \left\{(\mu,\sigma^2)|\mu = \mu_0 \text{ and } 0 \leq \sigma^2 < \infty\right\}$$
$$\Omega = \left\{(\mu,\sigma^2)|\mu \in \mathbb{R} \text{ and } 0 \leq \sigma^2 < \infty\right\}$$

- The likelihood function for the normal distribution is

$$L(\mu,\sigma^2) = \left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)^n \exp\left[-\frac{1}{2}\sum_{j=1}^{n}\left(\frac{y_j-\mu}{\sigma}\right)^2\right].$$

The one-sample t test is a GLRT (continued)

- It follows that

$$L(\omega_e) = \left[\frac{n/e}{2\pi \sum_{j=1}^{n} (y_j - \mu_0)^2} \right]^{n/2}$$

$$L(\Omega_e) = \left[\frac{n/e}{2\pi \sum_{j=1}^{n} (y_j - \overline{y})^2} \right]^{n/2}$$

- From the above, it follows that

$$\lambda = \left[1 + \frac{n(\overline{y} - \mu_0)^2}{\sum_{j=1}^{n} (y_j - \overline{y})^2} \right]^{-n/2} = \left(1 + \frac{t^2}{n-1} \right)^{-n/2}$$

  where we defined $t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}}$. Hence, as t^2 increases, λ decreases.
- Rejecting H_0 whenever λ too small means rejecting H_0 whenever t too large.
- But t is an observation of the r.v. T_{n-1}, so too large means $\geq t_{\alpha,n-1}$. $\square$

Two sample tests

Bruce M. Boghosian

- Sometimes, instead of comparing a sample mean to a (somehow) known mean, we wish to compare *two sample means*:
  - **Two sources:** Farm X and Farm Y each send 10 cases of barley. For both shipments, we quantify the quality of each case. We would like to compare μ_X to μ_Y. Note that this is different from comparing μ_X to a hypothesized μ_0. We might also wish to do hypothesis testing on $H_0 : \mu_X = \mu_Y$, etc.
  - **Two treatments:** Farm sends two shipments, X and Y, of barley, each consisting of 10 cases. We malt (soak in water) the barley of shipment X for 8 hours before roasting it over a peat fire, and that of shipment Y for 12 hours before roasting it over a peat fire. Then we quantify the quality of the malted and roasted barley in both cases, and we compare μ_X and μ_Y. Again, we might also wish to do hypothesis testing on $H_0 : \mu_X = \mu_Y$, etc.
- Likewise, we might also wish to compare *two sample variances*, e.g., $H_0 : \sigma_X^2 = \sigma_Y^2$

Testing $H_0 : \mu_X = \mu_Y$ when $\sigma_X = \sigma_Y$

Bruce M. Boghosian

- **Thm:**
  - Let $X_1, \ldots, X_n$ be the first random sample from $N(\mu_X, \sigma)$
  - Let $Y_1, \ldots, Y_n$ be the second random sample from $N(\mu_Y, \sigma)$
  - Let S_X^2 and S_Y^2 be the two sample variances
  - Let S_p^2 be the *pooled variance*,

$$S_p^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2} = \frac{\sum_{j=1}^n (X_j - \overline{X})^2 + \sum_{j=1}^m (Y_j - \overline{Y})^2}{n+m-2}$$

  - Then the quantity

$$T_{n+m-2} = \frac{(\overline{X} - \mu_x) - (\overline{Y} - \mu_Y)}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}}$$

  has a Student T distribution with $n+m-2$ df.

Testing $H_0 : \mu_X = \mu_Y$ when $\sigma_X = \sigma_Y$ (continued)

Bruce M. Boghosian

- **Pf:**
  - First note that we can write

$$T_{n+m-2} = \frac{(\overline{X} - \mu_x) - (\overline{Y} - \mu_Y)}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}} = \frac{\frac{\overline{X}-\overline{Y}-(\mu_X-\mu_Y)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}}{\sqrt{\frac{S_p^2}{\sigma^2}}}$$

$$= \frac{\frac{\overline{X}-\overline{Y}-(\mu_X-\mu_Y)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}}{\sqrt{\frac{1}{n+m-2}\left[\sum_{j=1}^n \left(\frac{X_j-\overline{X}}{\sigma}\right)^2 + \sum_{j=1}^m \left(\frac{Y_j-\overline{Y}}{\sigma}\right)^2\right]}}$$

  - $\mathrm{Var}\left(\overline{X} + \overline{Y}\right) = \frac{\sigma^2}{n} + \frac{\sigma^2}{m} = \sigma^2\left(\frac{1}{n} + \frac{1}{m}\right)$
  - So numerator of above is distributed as a standard normal

Testing $H_0 : \mu_X = \mu_Y$ when $\sigma_X = \sigma_Y$ (continued)

- **Pf:**
  - Turning our attention to the denominator of

$$T_{n+m-2} = \frac{\frac{\overline{X}-\overline{Y}-(\mu_X-\mu_Y)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}}{\sqrt{\frac{1}{n+m-2}\left[\sum_{j=1}^{n}\left(\frac{X_j-\overline{X}}{\sigma}\right)^2+\sum_{j=1}^{m}\left(\frac{Y_j-\overline{Y}}{\sigma}\right)^2\right]}}$$

  - We see that $\sum_{j=1}^{n}\left(\frac{X_j-\overline{X}}{\sigma}\right)^2$ and $\sum_{j=1}^{m}\left(\frac{Y_j-\overline{Y}}{\sigma}\right)^2$ are independent χ^2 r.v.s with $n-1$ and $m-1$ df, respectively.
  - Hence their sum U is χ^2 distributed with $n+m-2$ df.
  - Also the numerator and denominator above are independent.
  - Hence $T_{n+m-2} = \frac{Z}{\sqrt{\frac{U}{n+m-2}}}$ is Student-t distributed with $n+m-2$ df. $\square$

Testing $H_0 : \mu_X = \mu_Y$ when $\sigma_X = \sigma_Y$ (continued)

Bruce M. Boghosian

- **Thm.:** Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ be independent random samples from normal distributions with means μ_X and μ_Y, respectively, and with the same standard deviation σ.
- Since H_0 is $\mu_X = \mu_Y$, define the quantity $t = \frac{\overline{x}-\overline{y}}{s_p\sqrt{\frac{1}{n}+\frac{1}{m}}}$.
  - To test $H_0 : \mu_X = \mu_Y$ versus $H_1 : \mu_X > \mu_Y$ at the α level of significance, reject H_0 if $t \geq +t_{\alpha,n+m-2}$.
  - To test $H_0 : \mu_X = \mu_Y$ versus $H_1 : \mu_X < \mu_Y$ at the α level of significance, reject H_0 if $t \leq -t_{\alpha,n+m-2}$.
  - To test $H_0 : \mu_X = \mu_Y$ versus $H_1 : \mu_X \neq \mu_Y$ at the α level of significance, reject H_0 if either (a) $t \leq -t_{\alpha/2,n+m-2}$ or (b) $t \geq +t_{\alpha/2,n+m-2}$.

Example

- Were Mark Twain and Quintus Curtius Snodgrass the same person?
- Proportion of 3-letter words used in $n = 8$ writings of MT and $m = 10$ of QCS.

MT	0.225	0.262	0.217	0.240	0.230	0.229	0.235	0.217		
QCS	0.209	0.205	0.196	0.210	0.202	0.207	0.224	0.223	0.220	0.201

- We have $\overline{x} = \frac{1}{8}\sum_{j=1}^{8} x_j = 0.2319$ and $\overline{y} = \frac{1}{10}\sum_{j=1}^{10} y_j = 0.2097$. Is this close enough to conclude that $\mu_X = \mu_Y$?
- Set up hypothesis test with $H_0 : \mu_X = \mu_Y$ and $H_1 : \mu_X \neq \mu_Y$.
- Calculate $s_X^2 = 0.0002103$ and $s_Y = 0.0000955$
- Then calculate

$$s_p = \sqrt{\frac{(n-1)s_X^2 + (m-1)s_Y^2}{n+m-2}} = 0.0121.$$

Example (continued)

- Were Mark Twain and Quintus Curtius Snodgrass the same person?
- We have $\overline{x} = \frac{1}{8}\sum_{j=1}^{8} x_j = 0.2319$ and $\overline{y} = \frac{1}{10}\sum_{j=1}^{10} y_j = 0.2097$.
- Calculate $s_X^2 = 0.0002103$ and $s_Y = 0.0000955$
- Then calculate

$$s_p = \sqrt{\frac{(n-1)s_X^2 + (m-1)s_Y^2}{n+m-2}} = 0.0121.$$

$$t = \frac{\overline{x} - \overline{y}}{s_p\sqrt{\frac{1}{n} + \frac{1}{m}}} = \frac{0.2319 - 0.2097}{0.0121\sqrt{\frac{1}{8} + \frac{1}{10}}} = 3.88$$

- Take $\alpha = 0.01$, reject H_0 if $t \leq -t_{0.005,16} = -2.9208$ or $t \geq t_{0.005,16} = 2.9208$.
- Hence we reject H_0. MT & QCS not same person with 99% confidence.

The Behrens-Fisher problem

- Can we repeat the above analysis if $\sigma_X \neq \sigma_Y$?
- This is still an unsolved problem in statistics.
- Instead of $T = \frac{\overline{X}-\overline{Y}}{S_p\sqrt{\frac{1}{n}+\frac{1}{m}}}$, a widely used approximation is

$$W = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}}$$

- **Thm. (Welch 1938):** W is approximately distributed like a Student t distribution with

$$\frac{\left(\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}\right)^2}{\frac{\sigma_X^4}{n^2(n-1)} + \frac{\sigma_Y^4}{m^2(m-1)}},$$

rounded to the nearest integer.

Testing $H_0 : \sigma_X^2 = \sigma_Y^2$

- Let $x_1, \ldots, x_n$ and $y_1, \ldots, y_m$ be independent random samples from $N(\mu_X, \sigma_X)$ and $N(\mu_Y, \sigma_Y)$, respectively.
  - To test $H_0 : \sigma_X^2 = \sigma_Y^2$ versus $H_1 : \sigma_X^2 > \sigma_Y^2$ at the α level of significance, reject H_0 if $s_Y^2/s_X^2 \leq F_{\alpha,m-1,n-1}$.
  - To test $H_0 : \sigma_X^2 = \sigma_Y^2$ versus $H_1 : \sigma_X^2 < \sigma_Y^2$ at the α level of significance, reject H_0 if $s_Y^2/s_X^2 \geq F_{1-\alpha,m-1,n-1}$.
  - To test $H_0 : \sigma_X^2 = \sigma_Y^2$ versus $H_1 : \sigma_X^2 \neq \sigma_Y^2$ at the α level of significance, reject H_0 if either (a) $s_Y^2/s_X^2 \leq F_{\alpha/2,m-1,n-1}$ or (b) $s_Y^2/s_X^2 \geq F_{1-\alpha/2,m-1,n-1}$.

Example

- People with Raynaud's syndrome have impaired blood circulation to the fingers, causing heat loss.
- Measurements of heat output of fingers of $n = 10$ normal subjects, and $m = 10$ subjects with Raynaud's syndrome

x (wo/ RS)	2.43	1.83	2.43	2.70	1.88	1.96	1.53	2.08	1.85	2.44
y (wi/ RS)	0.81	0.70	0.74	0.36	0.75	0.56	0.65	0.87	0.40	0.31

- We have $\overline{x} = 2.11$, $s_X = 0.37$, and we have $\overline{y} = 0.62$ and $s_Y = 0.20$.
- It is evident that $\overline{Y} < \overline{X}$, but what about the variances?
- Test $H_0 : \sigma_X^2 = \sigma_Y^2$ versus $H_1 : \sigma_X^2 \neq \sigma_Y^2$ at the $\alpha = 0.05$ level of significance.
- Reject H_0 if either (a) $s_Y^2/s_X^2 \leq F_{\alpha/2,m-1,n-1}$ or (b) $s_Y^2/s_X^2 \geq F_{1-\alpha/2,m-1,n-1}$
- We have $F_{0.025,9,9} = 0.248$ and $F_{0.975,9,9} = 4.03$.
- Since ${}^{s_Y^2}/{}_{s_X^2} = 0.292$, we are unable to reject the null hypothesis that $\sigma_X^2 = \sigma_Y^2$.

Summary

Bruce M. Boghosian

- We have shown that the one-sample t test is a GLRT.
- We have defined two-sample tests.
- We have shown how to test $H_0 : \mu_X = \mu_Y$ and given an example.
- We have discussed the Behrens-Fisher problem, and presented an approximate solution, but we note that this is an important unsolved problem of statistics.
- We have shown how to test $H_0 : \sigma_X^2 = \sigma_Y^2$ and given an example.