Math 152. 4/23/09

Signed Rank Test

Last time, we saw that a paired sample design can reduce the variance of the mean of the differences (or difference in means in the unpaired case).

e.g.

Before	After	Diff	|Diff|	Rank	Signed Rank
25	27	2	2	2	2
29	25	−4	4	3	−3
60	59	−1	1	1	−1
27	37	10	10	4	4

$$W^+ = 2 + 4 = 6.$$

<u>Calculate differences.</u>
rank the abs. values of the differences.
restore the signs to the ranked abs values
sum the positive sign ranks. $= W^+$

Reject if W^+ is too large or too small.

under the null hypothesis:
(that the distribution of pair differences is symmetric about the origin).

each rank is equally likely to be + or − and each of

the 2^n (not all distinct sums)
combinations of signs and ranks. is equally likely to occur

The distribution of W^+ is tabulated for up to moderate sample sizes (and can be easily simulated).

For larger sample sizes ($n \geq 20$) the distribution is closely approximated by the normal distribution.

$$W^+ = \sum_{k=1}^{n} k\, I_k$$

where $I_k = \begin{cases} 1 & \text{if } k^{th} \text{ larges } |D_i| \text{ has } D_i > 0 \\ 0 & \text{else.} \end{cases}$

$$\text{Prob}\{I_k = 1\} = 1/2$$

under the Null

So $E(W^+) = \sum_{k=1}^{n} k \cdot \frac{1}{2} = \frac{1}{2}\,\frac{n(n+1)}{2}$

$= \frac{n(n+1)}{4}.$

$$\mathrm{Var}(W^+) = \sum_{k=1}^{n} k^2 \cdot \frac{1}{4}$$

$$= \frac{n(n+1)(2n+1)}{24}.$$

A more general C.L.T

$\Rightarrow$

$$\frac{W^+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \sim N(0,1).$$

$n \geq 20.$

Methods which assume normality

Un paired t-test.

"re call:" $\begin{cases} X_1, \ldots, X_n & \text{i.i.d. } N(\mu_x, \sigma^2) \\ Y_1, \ldots, Y_m & \text{i.i.d. } N(\mu_y, \sigma^2) \\ X_i, Y_j \text{ ind } \forall i j \end{cases}$

$$\Rightarrow \quad t = \frac{\bar{X} - \bar{Y} - (\mu_x - \mu_y)}{s_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{m+n-2}$$

where $S_p^2 = \dfrac{(n-1)S_X^2 + (m-1)S_Y^2}{m+n-2}$.

and t_{m+n-2} denotes the "t-distribution with $m+n-2$ degrees of freedom".

Def: if $Z \sim N(0,1)$, $U \sim \chi_n^2$ are indepent.

the distribution of

$Z / \sqrt{U/n}$ is called

the t-distribution with n degrees of freedom.

it can be shown that the density of $X \sim t_n$ is

$$f_X(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\,\Gamma(n/2)} \left(1 + \frac{t^2}{n}\right)^{-(\frac{n+1}{2})}$$

For $n \geq 20$ the t-distribution is approx. normal.

∴ a $100(1-\alpha)\%$ C.I for $\mu_X - \mu_Y$ is $\cancel{(\bar{X} - \bar{Y}) \pm t_{m+n-2}(\frac{\alpha}{2})}$

$$\bar{X} - \bar{Y} \pm t_{m+n-2}(\alpha/2)\, S_{(\bar{X} - \bar{Y})}$$

where $S_{\bar{X} - \bar{Y}} = S_p \sqrt{\frac{1}{n} + \frac{1}{m}}$

and $t_{m+n-2}(\frac{\alpha}{2}) = x$

s.t if

$X \sim t_{m+n-2}$ then $\text{Prob}\{X \geq x\} = \frac{\alpha}{2}$.

If the sample sizes are large the the t-test is robust against departures from normality. in the distribution of X and Y.

($\bar{X}$, $\bar{Y}$ are $\approx$ normal).

For smaller sample sizes
($\approx$ 20 or so)
the test is more robust than you might guess but can go wrong in unfavorable cases.

<u>typical problems:</u>

- 2 ppp's have same s.d's and similarly shaped distributions sample sizes roughly equal.
  $\Rightarrow$ long tails cause moderate trouble
  skewness does not (much)
- Same as above, but different sample sizes, [long tails + skewness = serious problem

- for two populations w/ considerably different skewness t-test can give very misleading results for small and moderate sample sizes.

- if independence assumptions are violated it is not wise to use "t-tools".

Two important/useful theoretical facts:

1) The two sided t-test:

$H_A : \mu_X \neq \mu_Y$) is equivalent to a likelihood ratio test.

2) if $Var(X) \neq Var(Y)$ then

$$Var(\bar{X} - \bar{Y}) \approx \frac{S_X^2}{n} + \frac{S_Y^2}{m}$$

is a natural substitute in the denominator of the t-statistic.

It can be shown that

$$df \approx \frac{\left[\left(S_X^2/n\right) + \left(S_Y^2/m\right)\right]^2}{\frac{\left(S_X^2/n\right)^2}{n-1} + \frac{\left(S_Y^2/m\right)^2}{m-1}}$$

to the nearest integer.

Paired Samples. (t-test)

if the differences between pairs are normally distributed with

$$E(D_i) = \mu_X - \mu_Y = \mu_D$$

and $Var(D_i) = \sigma_D^2$

(with unknown μ_D, σ_D)

then

$$t = \frac{\bar{D} - \mu_D}{S_{\bar{D}}}$$

follows a t-distribution with $n-1$ degrees of freedom.

A $100(1-\alpha)\%$ confidence interval for μ_D is

$$\bar{D} \pm t_{n-1}(\alpha/2)\, S_{\bar{D}}$$

(where $Prob\{X > t_{n-1}(\alpha/2)\} = \alpha/2$ when $X \sim t_{n-1}$).

The rejection criteria for a two sided test at level α of $H_0 : \mu_D = 0$ is

$$|\bar{D}| > t_{n-1}(\alpha/2)\, S_{\bar{D}}.$$

These results are approximately true in the absence of normality, for large sample sizes.

For small sample sizes and non-normal data, they can be in serious error.