STAT 824 sp 2025 Lec 12 slides

Wilcoxon rank-sum test

Karl Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Suppose we collect random samples from "control" and "treatment" populations:

$$X_1, \ldots, X_n \overset{\text{ind}}{\sim} F \qquad \text{"control"}$$

$$Y_1, \ldots, Y_m \overset{\text{ind}}{\sim} G \qquad \text{"treatment"}$$

We wish to test for treatment effectiveness (are Y's bigger than X's?).

Wilcoxon rank sum test (quintessential nonparametric test)

The *Wilcoxon rank sum test (WXRS)* concludes a "positive treatment effect" if

$$W_{XY} = \sum_{i=1}^{n} \sum_{j=1}^{m} \mathbf{1}(X_i \leq Y_j) \geq c,$$

where c can be calibrated to control the Type I error rate.

Can modify to find a "negative" or "either direction" treatment effect.

If $G = F$, the (null) distribution of W_{XY} is the same for any continuous F.

Exercise: For $X \sim F$ and $Y \sim G$, both continuous, show

1. $P(X < Y) = \int_0^1 F(G^{-1}(u))du$.
2. $P(X < Y) = 1/2$ if $F = G$.

Rank-sum form of Wilcoxon rank sum statistic

An alternate way of computing W_{XY}:

1. Sort all the data $(X_1, \ldots, X_n, Y_1, \ldots, Y_m)$
2. Obtain the ranks.
3. Keep the ranks corresponding to $Y_1, \ldots, Y_m$, calling these $R_1, \ldots, R_m$.

Then $W_{XY} = R_1 + \cdots + R_m - m(m+1)/2$.

Let $W_R = R_1 + \cdots + R_n$.

Exercise: Show that $W_{XY} = W_R - m(m+1)/2$.

Theorem

Let $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ be continuous iid rvs and set $N = n + m$. Then

$$P(\{R_1, \ldots, R_m\} = \{r_1, \ldots, r_m\}) = \frac{1}{\binom{N}{m}}$$

for all sets of m ranks $\{r_1, \ldots, r_m\} \subset \{1, \ldots, N\}$.

Exercise: Tabulate the null distribution of W_{XY} under $N = 5$, $m = 2$.

```
# generate some data
n <- 20
m <- 25
X <- rnorm(n,1,1)
Y <- rnorm(m,1,1)

# compute WR and WXY
Z <- c(X,Y)
id <- c(rep(1,n),rep(2,m))
R <- rank(Z)[id == 2]
WR <- sum(R)
WXY <- sum(R) - m*(m+1)/2

# must subtract 1, since we reject when WXY >= c
pval <- 1 - pwilcox(WXY-1,m = m, n = n)

# see that the wilcox.test() function gives the same values
wilcox.test(x=Y,y=X,alternative="greater",exact = TRUE) # switch X and Y
```

On computing the exact distribution of W_{XY} when N and n are large...

Theorem (Asymptotic Normality of rank sum under the null)

Under H_0: $F = G$ we have

$$\frac{W_R - \mathbb{E}W_R}{\sqrt{\text{Var } W_R}} \xrightarrow{d} \text{Normal}(0, 1)$$

as $n, m \to \infty$.

Exercise: Show $\mathbb{E}W_R = \frac{1}{2}m(N+1)$ and $\text{Var } W_R = \frac{1}{12}m(N-m)(N+1)$.

Corollary

An asymptotic p-value for testing H_0: $F = G$ versus the "right-sided" alternative is

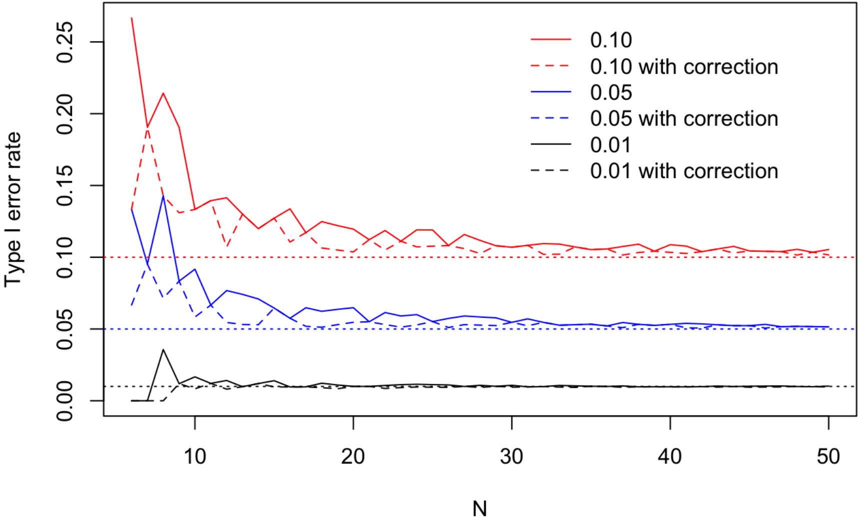
$$1 - \Phi\left(\frac{W_R - \frac{1}{2}m(N+1) - \frac{1}{2}}{\sqrt{\frac{1}{12}m(N-m)(N+1)}}\right),$$

where the extra $\frac{1}{2}$ is a "continuity correction".

Correspondingly, the rule which rejects H_0: $F = G$ when

$$W_{XY} \geq \mathbb{E}W_R + \frac{1}{2} - m(m+1)/2 + z_\alpha\sqrt{\text{Var } W_R}$$

has size approaching α as $n, m \to \infty$


0.10
0.10 with correction
0.05
0.05 with correction
0.01
0.01 with correction
Type I error rate
0.00
0.05
0.10
0.15
0.20
0.25
10
20
30
40
50
N

Sketch of asymptotic Normality proof

Assume H_0: $F = G$ and introduce $U_1, \dots, U_N \overset{\text{ind}}{\sim} \text{Uniform}(0,1)$. Then:

1. Write W_R as a sum of *dependent* rvs: $W_R = \sum_{i=1}^N i \cdot J_i, \quad J_i = \mathbf{1}(U_i \leq U_{(N)})$.
2. Introduce approximator $\tilde{W}_R$, which is a sum of *independent* rvs:

$$\tilde{W}_R = \sum_{i=1}^N \left(i - \frac{N+1}{2}\right) K_i + \frac{m(N+1)}{2}, \quad K_i = \mathbf{1}(U_i \leq m/N).$$

3. Show that $\dfrac{\tilde{W}_R - \mathbb{E}\tilde{W}_R}{\sqrt{\text{Var}\,\tilde{W}_R}} \xrightarrow{\text{d}} \text{Normal}(0,1)$ as $n, m \to \infty$.
4. Argue same holds for W_R since $\dfrac{\mathbb{E}(W_R - \tilde{W}_R)^2}{\text{Var}\,\tilde{W}_R} \to 0$ as $n, m \to \infty$.

Exercise:

1. Show $\mathbb{E}\tilde{W}_R = \mathbb{E}W_R$ and $\text{Var}\,\tilde{W}_R = \frac{(N-1)}{N}\text{Var}\,W_R$.
2. Show $\dfrac{\tilde{W}_R - \mathbb{E}\tilde{W}_R}{\sqrt{\text{Var}\,\tilde{W}_R}} \xrightarrow{\text{d}} \text{Normal}(0,1)$ as $n, m \to \infty$ with Lindeberg CLT.

To analyze the power of the WXRS we must specify an alternative to H_0: $F = G$.

Location shift model

In the *location-shift* model we assume $G(x) = F(x - \Delta)$ for some $\Delta \in \mathbb{R}$.

We will consider the right-sided test H_0: $\Delta \leq 0$ vs H_1: $\Delta > 0$.

Exercise: Show that the power of the rule $W_{XY} \geq c$ is nondecreasing in Δ.

It is convenient to use a Normal approximation to the power:

Theorem (Approximate power of WXRS in location-shift model)

In the location-shift model the power of $W_{XY} \geq c$ admits the approximation

$$\gamma(\Delta) \approx 1 - \Phi\left(\frac{c - nmp_1(\Delta)}{\sqrt{\vartheta(\Delta)}}\right)$$

provided N, n, and $N-n$ are all large, where $p_1(\Delta) = P(X_1 < Y_1)$ and

$$\vartheta(\Delta) = mnp_1(\Delta)[1-p_1(\Delta)] + mn(n-1)[p_2(\Delta) - p_1^2(\Delta)] + nm(m-1)[p_3(\Delta) - p_1^2(\Delta)]$$

with $p_2(\Delta) = P(X_1 < Y_1, X_2 < Y_1)$ and $p_3(\Delta) = P(X_1 < Y_1, X_1 < Y_2)$.

Exercise:

1. Establish the above result.
2. Find the value of c such that the test has size approximately equal to α.

Exercise: Show that making the substitutions

1. $c = c_\alpha$
2. $p_1(\Delta) = 1/2 + \Delta f^*(0)$, f^* the density of $X_1 - X_2$
3. $\vartheta(\Delta) = \vartheta(0)$

leads to the approximate power curve for the size-α test given by

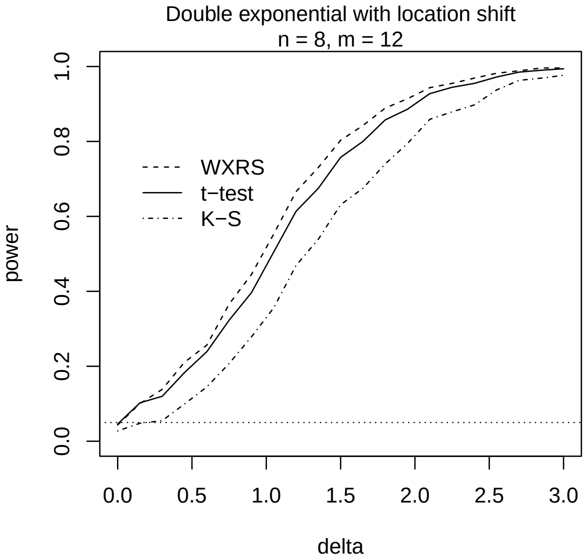
$$\tilde{\gamma}_\alpha(\Delta) = 1 - \Phi\left(z_\alpha - \sqrt{\frac{12nm}{N+1}} \cdot \Delta \cdot f^*(0)\right).$$

Exercise:

1. Show that if F is Normal, $n = m$, and $N + 1$ is replaced by $2n$, we obtain

$$\tilde{\gamma}_\alpha(\Delta) = 1 - \Phi\left(z_\alpha - \frac{\sqrt{6n}}{2\sigma\sqrt{\pi}} \cdot \Delta\right).$$

2. Find the smallest n such that the WXRS has power $\geq \gamma^*$ for all $\Delta \geq \Delta^*$.
3. Compare to n needed for the equal-variances two-sample z-test.


Double exponential with location shift
n = 8, m = 12
WXRS
t−test
K−S
power
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
delta