4. Two-sample Dispersion and Other Problems

4.1 Two-sample dispersion problem

- The "dispersion" of a random variable X (or its distribution) represents the variability of X, and can be measured by the variance of X.
- A large variance of X means that X can take values far away from its mean with high or moderate probability; whereas a small variance points to a high probability that X is close to its mean.
- In the extreme case of $\mathrm{Var}(X) = 0$, $\Pr(X = \mathrm{E}[X]) = 1$.
- In the two-sample dispersion problem, the data consist of two independent samples $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$, with a total of $N = m + n$ observations.
- The basic assumptions on $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ are the same as those in Assumption 3.1 (continuous and i.i.d. for each sample).
- We first consider the case that the only possible difference between the two samples lies in their variances.

Ansari-Bradley rank test for dispersion

Assumption 4.1 The cdf's $F(t)$ of $X_1,\ldots,X_m$ and $G(t)$ of $Y_1,\ldots,Y_n$ both satisfy the *location-scale parameter model* as follows:

$$F(t) = H\left(\frac{t-\theta}{\eta_1}\right) \quad \text{and} \quad G(t) = H\left(\frac{t-\theta}{\eta_2}\right), \quad t \in \mathbb{R}, \tag{4.1}$$

where $H(t)$ is a continuous cdf with median 0, $\theta \in \mathbb{R}$ is the *location parameter* and $\eta_1, \eta_2 > 0$ are the *scale parameters*.

Let $X \sim F(t)$ and $Y \sim G(t)$. Then model (4.1) can be equivalently expressed as

$$\frac{X-\theta}{\eta_1} \sim \frac{Y-\theta}{\eta_2} \tag{4.2}$$

Form (4.2), it is easy to see that $\mathrm{Var}(X) = \mathrm{Var}(Y) \Leftrightarrow \eta_1 = \eta_2$. Define

$$\gamma = \frac{\eta_1}{\eta_2} \quad \text{so that} \quad \gamma^2 = \frac{\mathrm{Var}(X)}{\mathrm{Var}(Y)} \tag{4.3}$$

Then $\mathrm{Var}(X) = \mathrm{Var}(Y) \Leftrightarrow \gamma^2 = 1$.

Null hypothesis: $H_0 : F(t) = G(t)$ for all $t \in \mathbb{R}$.

Under model (4.1), it is equivalent to $H_0 : \gamma^2 = 1$.

Alternative hypotheses:

$H_1 : \gamma^2 > 1$ ($\mathrm{Var}(X) > \mathrm{Var}(Y)$: X has more variability than Y);

$H_1 : \gamma^2 < 1$ ($\mathrm{Var}(X) < \mathrm{Var}(Y)$: X has less variability than Y); or

$H_1 : \gamma^2 \neq 1$ ($\mathrm{Var}(X) \neq \mathrm{Var}(Y)$: X and Y have different variability).

Rank scores: Assume no ties. Let $Z_1 < Z_2 < \cdots < Z_N$ be the ordered values of $X_1, \ldots, X_m, Y_1, \ldots, Y_n$ combined. Assign scores to $Z_1, \ldots, Z_N$ as follows:

- $1, 2, \ldots, \frac{N}{2} - 1, \frac{N}{2}, \frac{N}{2}, \frac{N}{2} - 1, \ldots, 2, 1$ if $N = m + n$ is even;
- $1, 2, \ldots, \frac{N-1}{2}, \frac{N+1}{2}, \frac{N-1}{2}, \ldots, 2, 1$ if N is odd.

(Score 1 to Z_1, Z_N, score 2 to Z_2, Z_{N-1}, and so on.)

Test statistic: Let R_j be the score of Y_j, $j = 1, \ldots, n$. Then the *Ansari-Bradley rank test* statistic for dispersion is defined by

$$C = \sum_{j=1}^{n} R_j \tag{4.4}$$

Rejection rule: Note that a large value of C means that more Y values tend to be in the middle range (less variability) than X, indicating $\mathrm{Var}(X) > \mathrm{Var}(Y)$, or $\gamma^2 > 1$. Hence the test rejects $H_0 : \gamma^2 = 1$ at the (achievable) α-level if

$C \geq c_\alpha$ for $H_1 : \gamma^2 > 1$, where $\Pr(C \geq c_\alpha) = \alpha$ with integer c_α;

$C \leq c_{1-\alpha} - 1$ for $H_1 : \gamma^2 < 1$ ($\Pr(C \leq c_{1-\alpha} - 1) = 1 - \Pr(C \geq c_{1-\alpha}) = 1 - (1-\alpha) = \alpha$);

either $C \geq c_{\alpha_1}$ or $C \leq c_{1-\alpha_2} - 1$ for $H_1 : \gamma^2 \neq 1$, where $\alpha_1 + \alpha_2 = \alpha$.

If $N = m + n$ is even, the distribution of C is symmetric. In this case, it is naturel to take $\alpha_1 = \alpha_2 = \alpha/2$. If N is odd, however, the distribution of C is asymmetric (not symmetric). An example will be shown later.

The distribution of C

Let $(a_1,\ldots,a_N)$ be the scores assigned to all ordered values of $X_1,\ldots,X_m,Y_1,\ldots,Y_n$. Then the possible scores $R_1,\ldots,R_n$ of $Y_1,\ldots,Y_n$ are given by

$$(R_1,\ldots,R_n)=(a_{i(1)},\ldots,a_{i(n)}) \text{ with } 1\le i(1)<i(2)<\cdots<i(n)\le N.$$

The total number of choices for $(R_1,\ldots,R_n)$ is thus $\binom{N}{n}$. The exact distribution of C under H_0 is then given by

$$\Pr(C=c)=\frac{\text{No. of } (R_1,\ldots,R_n): R_1+\cdots+R_n=c}{\binom{N}{n}} \tag{4.5}$$

If there are no ties, then

$$(a_1,\ldots,a_N)=\begin{cases}\left(1,2,\ldots,\frac{N}{2}-1,\frac{N}{2},\frac{N}{2},\frac{N}{2}-1,\ldots,2,1\right) & \text{for even } N=m+n\\ \left(1,2,\ldots,\frac{N-1}{2},\frac{N+1}{2},\frac{N-1}{2},\ldots,2,1\right) & \text{for odd } N=m+n\end{cases}$$

Example 4.1 Consider the case of $m = 4$ and $n = 3$. As $N = 4+3 = 7$ is odd,

$$(a_1, \ldots, a_7) = (1,2,3,4,3,2,1)$$

The total number of triplets $(R_1, R_2, R_3) = (a_i, a_j, a_k)$ with $1 \le i < j < k \le N$ is

$$\binom{N}{n} = \binom{7}{3} = \frac{7 \times 6 \times 5}{3 \times 2} = 7 \times 5 = 35$$

The smallest value of $C = R_1 + R_2 + R_3$ is 4 when $(R_1, R_2, R_3) = (a_1, a_2, a_7) = (1,2,1)$ and the largest value of C is 10 when $(R_1, R_2, R_3) = (a_3, a_4, a_5) = (3,4,3)$.

The range of C is $\{4,5,6,7,8,9,10\}$. The distribution of C under H_0 is worked out in Table 4.1 below. It is clearly asymmetric. From Table 4.1, we can obtain

$$\Pr(C \ge 10) = 1/35 = 0.029, \quad \Pr(C \ge 9) = 5/35 = 0.143, \quad \Pr(C \ge 8) = 12/35 = 0.343$$

$$\Pr(C \ge 7) = 20/35 = 0.571, \quad \Pr(C \ge 6) = 29/35 = 0.829, \quad \Pr(C \ge 5) = 33/35 = 0.943$$

Hence the achievable level includes $\alpha = 0.029$ with $c_\alpha = 10$ for $H_1 : \gamma^2 > 1$ and $\alpha = 0.057$ with $c_{1-\alpha} - 1 = c_{1-0.057} - 1 = c_{0.943} - 1 = 5 - 1 = 4$ for $H_1 : \gamma^2 < 1$.

Table 4.1 Distribution of C with $m = 4$ and $n = 3$ under H_0

c	$(R_1, \ldots, R_n) = (R_1, R_2, R_3) = (a_i, a_j, a_k),\ 1 \le i < j < k \le 7$	$\Pr(C = c)$
4	$(a_1, a_2, a_7) = (1,2,1),\ (a_1, a_6, a_7) = (1,2,1)$	$2/35$
5	$(a_1, a_2, a_6) = (1,2,2),\ (a_2, a_6, a_7) = (2,2,1),\ (a_1, a_3, a_7) = (1,3,1),$ $(a_1, a_5, a_7) = (1,3,1)$	$4/35$
6	$(a_1, a_2, a_3) = (1,2,3),\ (a_5, a_6, a_7) = (3,2,1),\ (a_1, a_3, a_6) = (1,3,2),$ $(a_2, a_5, a_7) = (2,3,1),\ (a_1, a_4, a_7) = (1,4,1),\ (a_2, a_3, a_7) = (2,3,1),$ $(a_1, a_5, a_6) = (1,3,2),\ (a_3, a_6, a_7) = (3,2,1),\ (a_1, a_2, a_5) = (1,2,3)$	$9/35$
7	$(a_1, a_2, a_4) = (1,2,4),\ (a_4, a_6, a_7) = (4,2,1),\ (a_1, a_3, a_5) = (1,3,3),$ $(a_3, a_5, a_7) = (3,3,1),\ (a_1, a_4, a_6) = (1,4,2),\ (a_2, a_4, a_7) = (2,4,1),$ $(a_2, a_3, a_6) = (2,3,2),\ (a_2, a_5, a_6) = (2,3,2)$	$8/35$
8	$(a_1, a_3, a_4) = (1,3,4),\ (a_4, a_5, a_7) = (4,3,1),\ (a_1, a_4, a_5) = (1,4,3),$ $(a_3, a_4, a_7) = (3,4,1),\ (a_2, a_3, a_5) = (2,3,3),\ (a_3, a_5, a_6) = (3,3,2),$ $(a_2, a_4, a_6) = (2,4,2)$	$7/35$
9	$(a_2, a_3, a_4) = (2,3,4),\ (a_4, a_5, a_6) = (4,3,2),\ (a_2, a_4, a_5) = (2,4,3),$ $(a_3, a_4, a_6) = (3,4,2)$	$4/35$
10	$(a_3, a_4, a_5) = (3,4,3)$	$1/35$

Mean and variance of C

Given $(a_1,\ldots,a_N)$ and X defined in Theorem 3.1, the mean and variance of C under H_0 can be derived from Theorem 3.1 as follows:

$$\mathrm{E}_0[C] = n\mathrm{E}[X] = n\frac{1}{N}\sum_{i=1}^{N} a_i \tag{4.6}$$

and

$$\mathrm{Var}_0(C) = \frac{mn}{N-1}\mathrm{Var}(X) = \frac{mn}{N-1}\left[\mathrm{E}\left[X^2\right] - \left(\mathrm{E}[X]\right)^2\right]$$

$$= \frac{mn}{N-1}\left[\frac{1}{N}\sum_{i=1}^{N} a_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} a_i\right)^2\right] \tag{4.7}$$

Assume no ties. If $N = m+n$ is even, then

$$(a_1,\ldots,a_N) = \left(1,2,\ldots,\frac{N}{2}-1,\frac{N}{2},\frac{N}{2},\frac{N}{2}-1,\ldots,2,1\right)$$

Hence

$$\frac{1}{N}\sum_{i=1}^{N} a_i = \frac{2}{N}\sum_{i=1}^{N/2} i = \frac{2}{N}\cdot\frac{1}{2}\cdot\frac{N}{2}\left(\frac{N}{2}+1\right) = \frac{N+2}{4} \tag{4.8}$$

and

$$\frac{1}{N}\sum_{i=1}^{N} a_i^2 = \frac{2}{N}\sum_{i=1}^{N/2} i^2 = \frac{2}{N}\cdot\frac{1}{6}\cdot\frac{N}{2}\left(\frac{N}{2}+1\right)(N+1) = \frac{(N+1)(N+2)}{12} \tag{4.9}$$

Substitute (4.8) and (4.9) into (4.6) and (4.7), respectively, we get

$$\mathrm{E}_0[C] = \frac{n}{N}\sum_{i=1}^{N} a_i = \frac{n(N+2)}{4} = \frac{n(m+n+2)}{4} \tag{4.10}$$

and

$$\mathrm{Var}_0(C) = \frac{mn}{N-1}\left[\frac{(N+1)(N+2)}{12} - \left(\frac{N+2}{4}\right)^2\right]$$

$$= \frac{mn(N+2)}{N-1}\cdot\frac{4(N+1)-3(N+2)}{48} = \frac{mn(N+2)(N-2)}{48(N-1)} \tag{4.11}$$

If $N = m + n$ is odd, then

$$(a_1, \ldots, a_N) = \left(1, 2, \ldots, \frac{N-1}{2}, \frac{N+1}{2}, \frac{N-1}{2}, \ldots, 2, 1\right)$$

Hence

$$\frac{1}{N}\sum_{i=1}^{N} a_i = \frac{2}{N}\sum_{i=1}^{(N-1)/2} i + \frac{1}{N}\cdot\frac{N+1}{2} = \frac{2}{N}\cdot\frac{1}{2}\cdot\frac{N-1}{2}\left(\frac{N-1}{2}+1\right) + \frac{N+1}{2N}$$

$$= \frac{(N-1)(N+1)}{4N} + \frac{N+1}{2N} = \frac{(N+1)(N-1+2)}{4N} = \frac{(N+1)^2}{4N} \qquad (4.12)$$

and

$$\frac{1}{N}\sum_{i=1}^{N} a_i^2 = \frac{2}{N}\sum_{i=1}^{(N-1)/2} i^2 + \frac{1}{N}\left(\frac{N+1}{2}\right)^2 = \frac{N-1}{6N}\left(\frac{N-1}{2}+1\right)(N-1+1) + \frac{(N+1)^2}{4N}$$

$$= \frac{(N-1)(N+1)N}{12N} + \frac{(N+1)^2}{4N} = \frac{(N+1)\left(N^2+2N+3\right)}{12N} \qquad (4.13)$$

It follows from (4.12) and (4.13) that

$$\frac{1}{N}\sum_{i=1}^{N} a_i^2 - \left(\frac{1}{N}\sum_{i=1}^{N} a_i\right)^2 = \frac{(N+1)(N^2+2N+3)}{12N} - \frac{(N+1)^4}{16N^2}$$

$$= \frac{(N+1)}{48N^2}\left[4(N^3+2N^2+3N) - 3(N^3+3N^2+3N+1)\right]$$

$$= \frac{(N+1)(N^3-N^2+3N-3)}{48N^2} = \frac{(N+1)(N^2+3)(N-1)}{48N^2} \qquad (4.14)$$

Substitute (4.12) and (4.14) into (4.6) and (4.7), respectively, to obtain

$$\mathrm{E}_0[C] = n\frac{1}{N}\sum_{i=1}^{N} a_i = \frac{n(N+1)^2}{4N} \qquad (4.15)$$

and

$$\mathrm{Var}_0(C) = \frac{mn(N+1)(N^2+3)}{48N^2} \qquad (4.16)$$

Example 4.2 Let $m = 4$, $n = 3$, and $N = 4+3 = 7$. The mean and variance of C under H_0 can be calculated directly from the distribution in Table 4.1:

$$\mathrm{E}_0[C] = \frac{4(2)+5(4)+6(9)+7(8)+8(7)+9(4)+10}{35} = \frac{240}{35} = \frac{48}{7}$$

and

$$\mathrm{Var}_0(C) = \frac{4^2(2)+5^2(4)+6^2(9)+7^2(8)+8^2(7)+9^2(4)+10^2}{35} - \left(\frac{48}{7}\right)^2 = \frac{104}{49}$$

But it is much easier to do so by formulae (4.15) and (4.16):

$$\mathrm{E}_0[C] = \frac{n(N+1)^2}{4N} = \frac{3(7+1)^2}{4\times 7} = \frac{192}{28} = \frac{48}{7}$$

and

$$\mathrm{Var}_0(C) = \frac{mn(N+1)(N^2+3)}{48N^2} = \frac{4\times 3(7+1)(7^2+3)}{48\times 7^2} = \frac{96\times 52}{48\times 49} = \frac{104}{49}$$

Approximate distribution of C

Under $H_0: \gamma^2 = 1$, if n is large, then

$$C^* = \frac{C - \mathrm{E}_0[C]}{\sqrt{\mathrm{Var}_0(C)}} \sim N(0,1) \text{ approximately,} \qquad (4.17)$$

where $\mathrm{E}_0[C]$ and $\mathrm{Var}_0(C)$ are given by (4.10) – (4.11) for even $N = m + n$, or by (4.15) – (4.16) for odd N.

Approximate rejection rule: Reject $H_0: \gamma^2 = 1$ at the α level if

- $C^* \geq z_\alpha$ against $H_1: \gamma^2 > 1$;
- $C^* \leq -z_\alpha$ against $H_1: \gamma^2 < 1$;
- $|C^*| \geq z_{\alpha/2}$ against $H_1: \gamma^2 \neq 1$,

where C^* is defined in (4.17).

The *p*-value of the Ansari-Bradley test

Let c denote the observed value of C from the data. If $N = m + n$ is even, then C is symmetric about $a = \mathrm{E}_0[C]$, hence its p-value against the two-sided alternative $H_1 : \gamma^2 \neq 1$ is $\Pr(C \geq c) + \Pr(C \leq 2a - c) = 2\Pr(C \geq c)$ if $c > a$ or $2\Pr(C \leq c)$ if $c < a$ as for many symmetric test statistics.

If $N = m + n$ is odd, then C is not symmetric, hence the above rule does not apply. In that case, we define $p\text{-value} = 2\min\{\Pr(C \geq c), \Pr(C \leq c)\}$ for $H_1 : \gamma^2 \neq 1$, which still represents the probability of "more extreme than observed". This definition is valid for symmetric C as well because $\min\{\Pr(C \geq c), \Pr(C \leq c)\} = \Pr(C \geq c)$ if $c > a$, or $\Pr(C \leq c)$ if $c < a$.

The p-value for one-sided $H_1 : \gamma^2 > 1$ ($\gamma^2 < 1$) remains $\Pr(C \geq c)$ ($\Pr(C \leq c)$).

Moreover, if b is the observed sum of rank scores of X values, then $b + c = TS$, so that $c = TS - b$, where TS is the total score of all X and Y values.

By (4.8) and (4.12), $TS = N(N+2)/4$ for even N and $TS = (N+1)^2/4$ for odd N.

Ties: Like in the location problems, if there are ties among $X_1,\ldots,X_m,Y_1,\ldots,Y_n$, then average ranks will be assigned to tied values.

The exact distribution of C conditional on ties under H_0 can be worked out by (4.5) with scores adjusted for ties.

Mean and variance with ties

Let $r_1,\ldots,r_N$ denote the scores assigned to all $X_1,\ldots,X_m,Y_1,\ldots,Y_n$, with average scores for tied values.

Take $(a_1,\ldots,a_N)=(r_1,\ldots,r_N)$. Then the sum

$$\sum_{i=1}^{N} a_i = \sum_{i=1}^{N} r_i$$

is not affected by average ranks for ties. Hence the formulae for the mean $\mathrm{E}_0[C]$ under H_0 in (4.10) and (4.15) remain unchanged.

The variance $\mathrm{Var}_0(C)$, however, will be affected by average ranks for ties.

Since the sum of squares

$$\sum_{i=1}^{N} a_i^2 = \sum_{i=1}^{N} r_i^2$$

differs between the cases with and without ties, formulae (4.11) and (4.16) are no longer valid for $\mathrm{Var}_0(C)$. But formulae (4.6) and (4.7) remain valid with scores $(a_1,\ldots,a_N) = (r_1,\ldots,r_N)$ adjusted for ties.

Consequently, by substituting (4.8) and (4.12) into (4.6) and (4.7), we obtain the following formulae for $\mathrm{Var}_0(C)$, which are valid with or without ties:

$$\mathrm{Var}_0(C) = \frac{mn}{N(N-1)}\left[\sum_{j=1}^{N} r_j^2 - \frac{N(N+2)^2}{16}\right] \quad \text{if } N \text{ is even;} \qquad (4.18)$$

or

$$\mathrm{Var}_0(C) = \frac{mn}{N(N-1)}\left[\sum_{j=1}^{N} r_j^2 - \frac{(N+1)^4}{16N}\right] \quad \text{if } N \text{ is odd,} \qquad (4.19)$$

Example 4.3 Let $(X_1, X_2, X_3, X_4) = (2,2,4,4)$ and $(Y_1, Y_2, Y_3) = (1,2,8)$. Then $(Z_1, \ldots, Z_7) = (Y_1, X_1, X_2, Y_2, X_3, X_4, Y_3) = (1,2,2,2,4,4,8)$. Hence the scores for the combined values $(Z_1, \ldots, Z_7)$ take averages for $r_2 = r_3 = r_4$ and $r_5 = r_6$ from $(1,2,3,4,3,2,1)$, resulting in $(r_1, \ldots, r_7) = (1,3,3,3,2.5,2.5,1)$.

The mean of C conditional on ties under H_0 is $\mathrm{E}_0[C] = 48/7$ as in Example 4.2. But the variance, by (4.19), is reduced from $104/49$ in Example 4.2 (no ties) to

$$\mathrm{Var}_0(C) = \frac{4 \times 3}{7(7-1)} \left[2 \times 1^2 + 3 \times 3^2 + 2 \times 2.5^2 - \frac{(7+1)^4}{16 \times 7} \right] = \frac{69}{49}$$

Since $(Y_1, Y_2, Y_3) = (Z_1, Z_4, Z_7)$, $C = r_1 + r_4 + r_7 = 1 + 3 + 1 = 5$ and so

$$C^* = \frac{C - \mathrm{E}_0[C]}{\sqrt{\mathrm{Var}_0(C)}} = \frac{5 - 48/7}{\sqrt{69/49}} = -1.565$$

Thus the approximate p-value against $H_1 : \gamma^2 < 1$ is $\Pr(C^* \le -1.565) \approx 0.0588$. This shows that the achieved level of significance for $\gamma^2 < 1$ is about 5.88%.

Example 4.4 Refer to Example 5.1 of the textbook (on page 157). The data are Serum Iron (μg/100ml) Determination from two techniques: a new Jung-Parekh method (X) and an old Ramsay method (Y). One question of concern is whether there is a loss of accuracy when the Jung-Parekh method is used instead of the Ramsay method. This corresponds to $\text{Var}(X) > \text{Var}(Y)$, or $\gamma^2 > 1$.

The data are presented in Table 5.1 of the textbook (page 157). The scores of the combined values from X and Y are listed below:

Y	X	Y	Y	X	Y	Y	X	X	Y	Y	Y	Y	X	(14)
96	96	97	98	98	99	99	99	100	100	101	102	102	103	
1.5	1.5	3	4.5	4.5	7	7	7	9.5	9.5	11	12.5	12.5	14.5	
X	X	Y	X	X	X	X	X	Y	Y	X	Y	Y		(27)
103	104	104	104	105	105	106	106	106	107	107	107	108		
14.5	17	17	17	19.5	19.5	19	19	19	16	16	16	12.5		
X	X	X	Y	X	Y	Y	X	Y	Y	X	X	Y		(40)
108	108	108	109	110	110	111	113	113	113	114	114	116		
12.5	12.5	12.5	10	8.5	8.5	7	5	5	5	2.5	2.5	1		

Thus $C = 1.5 + 3 + 4.5 + 2 \times 7 + \cdots + 2 \times 5 + 1 = 185.5$ and

$$\sum_{j=1}^{n} r_j^2 = 2 \times 1.5^2 + 3^2 + 2 \times 4.5^2 + 3 \times 7^2 + \cdots + 2 \times 2.5^2 + 1^2 = 5721$$

Since $m = n = 20 \Rightarrow N = 40$, $\mathrm{E}_0[C] = 20(40+2)/4 = 210$ and by (4.18),

$$\mathrm{Var}_0(C) = \frac{20 \times 20}{40(40-1)} \left[5721 - \frac{40(40+2)^2}{16} \right] = \frac{13110}{39} = 336.15$$

Thus

$$C^* = \frac{C - \mathrm{E}_0[C]}{\sqrt{\mathrm{Var}_0(C)}} = \frac{185.5 - 210}{\sqrt{336.15}} = -1.336$$

This shows no evidence for $\gamma^2 > 1$ (the p-value > 0.5 for $\gamma^2 > 1$ if $C^* < 0$).

Instead, the p-value for $\gamma^2 < 1$ is $\Pr(C^* \le -1.336) \approx 0.0908 < 0.1$. Hence there is moderate evidence at the 10% level that $\gamma^2 < 1$, or $\mathrm{Var}(X) < \mathrm{Var}(Y)$, indicating that the Jung-Parekh method is actually more accurate than the Ramsay method.

Miller's Jackknife test

Jackknife estimation

The idea of Jackknife estimation is to estimate a quantity from each subsample omitting one observation from a sample of size n, which produces n estimates. Then we can use these estimates to estimate the variance, bias and mean-squared error of the estimator.

For example, given a sample $x_1,\ldots,x_n$, the estimate of the population mean μ by omitting the i-th observation is

$$\overline{x}_i = \frac{1}{n-1}\sum_{j=1,j\neq i}^{n} x_j, \quad i=1,\ldots,n.$$

The average of $\overline{x}_1,\ldots,\overline{x}_n$ is equal to the sample mean $\overline{x}$:

$$\frac{1}{n}\sum_{i=1}^{n}\overline{x}_i = \frac{1}{n(n-1)}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} x_j = \frac{1}{n(n-1)}\sum_{i=1}^{n}(n-1)x_i = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x}$$

Then a Jackknife estimate of $\mathrm{Var}(\bar{X})$ is given by

$$\widehat{\mathrm{Var}}(\bar{X}) = \frac{n-1}{n}\sum_{i=1}^{n}(\bar{x}_i - \bar{x})^2 = \frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

The Jackknife method can be applied to reduce the bias of a biased estimate. Note that although $\bar{X}$ is unbiased for μ, the MLE $\hat{\theta}$ of a quantity θ is generally biased, including the MLE of the variance σ^2 for the normal distribution.

Let $\hat{\theta}$ be an estimate of θ based on $x_1,\ldots,x_n$, $\hat{\theta}_{(i)}$ the estimate of θ by omitting x_i and $\hat{\theta}_{(\cdot)}$ the average of $\hat{\theta}_{(1)},\ldots,\hat{\theta}_{(n)}$. The Jackknife estimate of the bias of $\hat{\theta}$ for θ is $\widehat{\mathrm{Bias}}_{\hat{\theta}}(\theta) = (n-1)\left(\hat{\theta}_{(\cdot)} - \hat{\theta}\right)$, and the bias-corrected Jackknife estimate of θ is

$$\hat{\theta}_{\mathrm{J}} = \hat{\theta} - (n-1)\left(\hat{\theta}_{(\cdot)} - \hat{\theta}\right) = n\hat{\theta} - (n-1)\hat{\theta}_{(\cdot)} = \frac{1}{n}\sum_{i=1}^{n}\left[n\hat{\theta} - (n-1)\hat{\theta}_{(i)}\right]$$

Statistical theory has shown that $\hat{\theta}_{\mathrm{J}}$ can generally reduce the bias of an estimate $\hat{\theta}$ from the order of n^{-1} to n^{-2}.

Now we introduce the Miller's Jackknife test for dispersion.

Assumption 4.2

- The cdf's $F(t)$ of $X_1,\ldots,X_m$ and $G(t)$ of $Y_1,\ldots,Y_n$ both have the forms under location-scale parameter model:

$$F(t) = H\left(\frac{t-\theta_1}{\eta_1}\right) \quad \text{and} \quad G(t) = H\left(\frac{t-\theta_2}{\eta_2}\right), \quad t \in \mathbb{R}, \tag{4.20}$$

  where $H(t)$ is a continuous cdf with median 0, $\theta_1, \theta_2 \in \mathbb{R}$ are the location parameters, and $\eta_1, \eta_2 > 0$ the scale parameters.

- $H(t)$ has a finite 4th moment: $\mathrm{E}[V^4] < \infty$ for $V \sim H(t)$.

Note that (4.20) does not require $\theta_1 = \theta_2$.

Null hypothesis: $H_0 : \gamma^2 = 1$

Alternative hypotheses: $H_1 : \gamma^2 > 1$, $\gamma^2 < 1$, or $\gamma^2 \neq 1$.

Motivation: The idea of the Miller's Jackknife test is to reject $H_0 : \gamma^2 = 1$ for $H_1 : \gamma^2 = \mathrm{Var}(X)/\mathrm{Var}(Y) > 1$ if $\widehat{\mathrm{Var}}(X)/\widehat{\mathrm{Var}}(Y)$ is large, or equivalently, if

$$\log \frac{\widehat{\mathrm{Var}}(X)}{\widehat{\mathrm{Var}}(Y)} = \log \widehat{\mathrm{Var}}(X) - \log \widehat{\mathrm{Var}}(Y) \quad \text{is large}$$

based on the Jackknife estimates $\widehat{\mathrm{Var}}(X)$ of $\mathrm{Var}(X)$ and $\widehat{\mathrm{Var}}(Y)$ of $\mathrm{Var}(Y)$.

Test statistic: Assume $m > 2$ and $n > 2$. Define

- $\bar{X}_0 = \frac{1}{m}\sum_{i=1}^{m} X_i$, $D_0^2 = \frac{1}{m-1}\sum_{i=1}^{m}(X_i - \bar{X}_0)^2$;
- $\bar{Y}_0 = \frac{1}{n}\sum_{j=1}^{n} Y_j$, $E_0^2 = \frac{1}{n-1}\sum_{j=1}^{n}(Y_j - \bar{Y}_0)^2$;
- $\bar{X}_i = \frac{1}{m-1}\sum_{s\neq i}^{m} X_s$, $D_i^2 = \frac{1}{m-2}\sum_{s\neq i}^{m}(X_s - \bar{X}_i)^2$, $i = 1,\ldots,m$;

- $\bar{Y}_j = \frac{1}{n-1}\sum_{t\neq j}^{n} Y_t, \quad E_j^2 = \frac{1}{n-2}\sum_{y\neq j}^{n}(Y_t - \bar{Y}_j)^2, \quad j = 1,\ldots,n;$
- $S_i = \log D_i^2, \ i = 0,1,\ldots,m; \quad T_j = \log E_j^2, \ j = 0,1,\ldots,n.$

Then S_0 is an estimate of $\log(\mathrm{Var}(X))$ and S_i is an estimate of $\log(\mathrm{Var}(X))$ with missing X_i. Hence the Jackknife estimate of $\log(\mathrm{Var}(X))$ is given by

$$\bar{A} = \frac{1}{m}\sum_{i=1}^{m} A_i = \frac{1}{m}\sum_{i=1}^{m}\left[mS_0 - (m-1)S_i\right] = mS_0 - (m-1)S_{(\cdot)},$$

where $A_i = mS_0 - (m-1)S_i, \ i = 1,\ldots,m.$

Similarly, the Jackknife estimate of $\log(\mathrm{Var}(Y))$ is

$$\bar{B} = \frac{1}{n}\sum_{j=1}^{n} B_j = \frac{1}{n}\sum_{j=1}^{n}\left[nT_0 - (n-1)T_j\right] = nT_0 - (n-1)T_{(\cdot)},$$

where $B_j = nT_0 - (n-1)T_j, \ j = 1,\ldots,n.$

The Jackknife estimates of the variances of $\overline{A}$ and $\overline{B}$ based on sample variances of $A_1,\ldots,A_m$ and $B_1,\ldots,B_n$ are given by

$$V_1 = \sum_{i=1}^{m} \frac{(A_i - \overline{A})^2}{m(m-1)} \quad \text{and} \quad V_2 = \sum_{j=1}^{n} \frac{(B_i - \overline{B})^2}{n(n-1)}$$

The Miller's Jackknife test rejects $H_0 : \gamma^2 = 1$ in favour of $H_1 : \gamma^2 > 1$ if $\overline{A} - \overline{B}$ is large. Hence the test statistic is defined by

$$Q = \frac{\overline{A} - \overline{B}}{\sqrt{V_1 + V_2}} \sim N(0,1) \text{ approximately} \qquad (4.21)$$

Approximate rejection rule: Reject $H_0 : \gamma^2 = 1$ at the α level if

- $Q \geq z_\alpha$ against $H_1 : \gamma^2 > 1$;
- $Q \leq -z_\alpha$ against $H_1 : \gamma^2 < 1$;
- $|Q| \geq z_{\alpha/2}$ against $H_1 : \gamma^2 \neq 1$, where Q is defined in (4.21).

Example 4.5 Example 5.2 of the textbook (page 172) provide two-sample data:

$$(X_1,\ldots,X_5)=(6.2,5.9,8.9,6.5,8.6) \quad \text{and} \quad (Y_1,\ldots,Y_5)=(9.5,9.8,9.5,9.6,10.3)$$

The values of *Y*'s are clearly greater than those of *X*'s, hence it is not sensible to directly apply the Ansari-Bradley rank test to assess the difference in dispersion between *X* and *Y*. But this is not a problem for the Millar's Jackknife test.

Calculate

$$\bar{X}_0=\frac{6.2+5.9+8.9+6.5+8.6}{5}=7.22, \quad \bar{X}_1=\frac{5.9+8.9+6.5+8.6}{4}=7.745$$

$$D_0^2=\frac{(6.2-7.22)^2+(5.9-7.22)^2+\cdots+(8.6-7.22)^2}{4}=2.007$$

$$D_1^2=\frac{(5.9-7.745)^2+(8.9-7.745)^2+(6.5-7.745)^2+(8.6-7.745)^2}{3}=2.2425$$

Similarly, $\bar{X}_2=7.55$, $\bar{X}_3=6.8$, $\bar{X}_4=7.4$, $\bar{X}_5=6.875$;

$D_2^2=1.95$, $D_3^2=1.5$, $D_4^2=2.46$, $D_5^2=1.8825$;

$\bar{Y}_0 = 9.74,\ \bar{Y}_1 = 9.8,\ \bar{Y}_2 = 9.725,\ \bar{Y}_3 = 9.8,\ \bar{Y}_4 = 9.775,\ \bar{Y}_5 = 9.6;$

$E_0^2 = 0.113,\ E_1^2 = 0.1267,\ E_2^2 = 0.1492,\ E_3^2 = 0.1267,\ E_4^2 = 0.1425,\ E_5^2 = 0.02.$

It follows that $A_1 = 5\log D_0^2 - 4\log D_1^2 = 5(0.6966) - 4(0.8076) = 0.2526$

Similarly, $A_2 = 0.8118,\ A_3 = 1.861,\ A_4 = -0.1178,\ A_5 = 0.9526;$

$B_1 = -2.6372,\ B_2 = -3.2912,\ B_3 = -2.6372,\ B_4 = -3.1084,\ B_5 = 4.746$

$$\Rightarrow \begin{matrix} \bar{A} = 0.7520 \\ \bar{B} = -1.3856 \end{matrix}, \quad V_1 = \sum_{i=1}^{5} \frac{(A_i - \bar{A})^2}{5(5-1)} = 0.1140, \quad V_2 = \sum_{j=1}^{5} \frac{(B_i - \bar{B})^2}{5(5-1)} = 2.3664$$

$$\Rightarrow Q = \frac{\bar{A} - \bar{B}}{\sqrt{V_1 + V_2}} = \frac{0.7520 - (-1.3856)}{\sqrt{0.114 + 2.3664}} = 1.36 \quad \text{by (4.21)}$$

Thus the p-value for $\gamma^2 > 1$ by Miller's Jackknife test is $\Pr(Q \geq 1.36) \approx 0.0869$.
This shows moderate evidence that X has greater variability (dispersion) than Y.
Refer to Example 5.2 of the textbook for more details.

4.2 Location-dispersion problem

Assume the cdf's $F(t)$ of $X_1,\ldots,X_m$ and $G(t)$ of $Y_1,\ldots,Y_n$ to satisfy (4.20). The *location-dispersion* problem is to test whether there is a difference between $F(t)$ and $G(t)$, in either location or dispersion, or both.

Lepage rank test for either location or dispersion

This test combines the rank test statistics W for location and C for dispersion.

Hypotheses:

$H_0 : F(t) = G(t)$ for all $t \in \mathbb{R}$, or equivalently, $H_0 : \theta_1 = \theta_2$ and $\eta_1 = \eta_2$.

H_1 : Either $\theta_1 \neq \theta_2$ or $\eta_1 \neq \eta_2$ (or both).

Test statistic:

$$D = (W^*)^2 + (C^*)^2 = \frac{(W - \mathrm{E}_0[W])^2}{\mathrm{Var}_0(W)} + \frac{(C - \mathrm{E}_0[C])^2}{\mathrm{Var}_0(C)}, \qquad (4.22)$$

where W and C are defined in (3.4) and (4.4), respectively.

The distribution of D: By (4.22), the exact distribution of D under H_0 can be obtained by combining the distributions of W and C under H_0 as follows:

$$\Pr(D = d) = \frac{\text{No. of } (S_1, \ldots, S_n, R_1, \ldots, R_n) : (w^*)^2 + (c^*)^2 = d}{\binom{N}{n}} \qquad (4.23)$$

where $S_1, \ldots, S_n$ and $R_1, \ldots, R_n$ are respectively the ranks and scores for W and C,

$$w^* = \frac{S_1 + \cdots + S_n - \mathrm{E}_0[W]}{\sqrt{\mathrm{Var}_0(W)}} \quad \text{and} \quad c^* = \frac{R_1 + \cdots + R_n - \mathrm{E}_0[C]}{\sqrt{\mathrm{Var}_0(C)}}$$

Rejection rule: Reject H_0 at the α-level if $D \geq d_\alpha$, where d_α is a value of D such that $\Pr(D \geq d_\alpha) = \alpha$ under H_0, which is determined by (4.23).

Approximate rejection rule: Reject H_0 at the α-level if $D \geq \chi^2_{2,\alpha}$, where $\chi^2_{2,\alpha}$ is the upper α-percentile of the chi-square distribution χ^2_2 with two degrees of freedom, i.e., $\Pr(Q \geq \chi^2_{2,\alpha}) = \alpha$ with $Q \sim \chi^2_2$.

Example 4.6 In Example 5.3 of the textbook (page 183), platelet counts were recorded on two groups of newborn infants: one group (Y) with mothers treated by the steroid prednisone, and the other is a control group (X) (no treatment).

The problem is to determine the evidence of difference between the two groups in either location or dispersion (or both).

The data (in 000's per mm) are combined and ordered as follows, along with their ranks and Ansari-Bradley scores:

	X	X	X	X	X	Y	Y	Y	X	Y	Y	Y	Y	Y	Y	Y
	12	20	32	40	60	67	90	95	112	120	124	135	180	190	215	399
Rank	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Score	1	2	3	4	5	6	7	8	8	7	6	5	4	3	2	1

Thus the Wilcoxon rank sum W and the Ansari-Bradley statistic C are given by

$W = 6+7+8+10+11+12+13+14+15+16 = 112$ and

$C = 6+7+8+7+6+5+4+3+2+1 = 49$

Furthermore, since $m = 6$, $n = 10$ and $N = 16$, the means and variances of W and C under the null hypothesis of no difference in location or dispersion are

$$\mathrm{E}_0[W] = \frac{10(16+1)}{2} = 85, \quad \mathrm{Var}_0(W) = \frac{6 \times 10(16+1)}{12} = 85$$

and

$$\mathrm{E}_0[C] = \frac{10(16+2)}{4} = 45, \quad \mathrm{Var}_0(C) = \frac{6 \times 10(16+2)(16-2)}{48(16-1)} = 21$$

Thus the Lepage rank test statistic for location or dispersion is

$$D = \frac{(112-85)^2}{85} + \frac{(49-45)^2}{21} = 8.576 + 0.762 = 9.338$$

The exact p-value of the Lepage rank test is $\Pr(D \geq 9.338) = 0.0035$ by R, and the approximate p-value is $\Pr(\chi_2^2 \geq 9.338) = 0.0094$. These show very strong evidence that X and Y have different location or dispersion, indicating that the steroid therapy has a significant effect on the platelet counts of newborn infants.

Remark 4.1

In practice, after the Lepage rank test rejects $H_0 : \theta_1 = \theta_2$ and $\eta_1 = \eta_2$, the next step is to identify which difference is significant (location or dispersion or both).

In Example 4.6, the value of D is mainly contributed by W, whereas C is near 0. Hence it may be intuitively tempting to think $\theta_1 \neq \theta_2$ and $\eta_1 = \eta_2$ as reasonable. Such an intuition, however, may not be right, because the Ansari-Bradley test requires $\theta_1 = \theta_2$ and it may be unable to detect the difference in dispersion if θ_1 and θ_2 differ significantly. On the other hand, the Wilcoxon rank sum test requires condition (3.1) or (3.2), which implies $\mathrm{Var}(X) = \mathrm{Var}(Y)$ or $\eta_1 = \eta_2$. Thus a large value of W does not necessarily indicate $\theta_1 \neq \theta_2$ if $\eta_1 \neq \eta_2$.

To check such an intuition statistically, a sensible approach is to first carry out a Miller's Jackknife test, which does not require $\theta_1 = \theta_2$. If it accepts $H_0 : \eta_1 = \eta_2$, then the Wilcoxon rank sum test is appropriate for $H_0 : \theta_1 = \theta_2$. If the Miller's Jackknife test rejects $\eta_1 = \eta_2$, then this intuition is not justified.

4.3 General differences in two populations

Now consider the problem of testing the difference between $F(t)$ of $X_1,\ldots,X_m$ and $G(t)$ of $Y_1,\ldots,Y_n$ just based on basic assumptions, without requiring such assumptions as the location-dispersion parameter model in (4.20).

Two-sample Kolmogorov-Smirnov test

Hypothesis: $H_0 : F(t) = G(t),\ t \in \mathbb{R}$, against $H_1 : F(t) \neq G(t)$ for some $t \in \mathbb{R}$.

Test statistic: Let $d =$ greatest common divisor of m and n. Denote by

$$F_m(t) = \frac{1}{m}\sum_{i=1}^{m} I_{\{X_i \le t\}} \quad \text{and} \quad G_n(t) = \frac{1}{n}\sum_{j=1}^{n} I_{\{Y_j \le t\}}$$

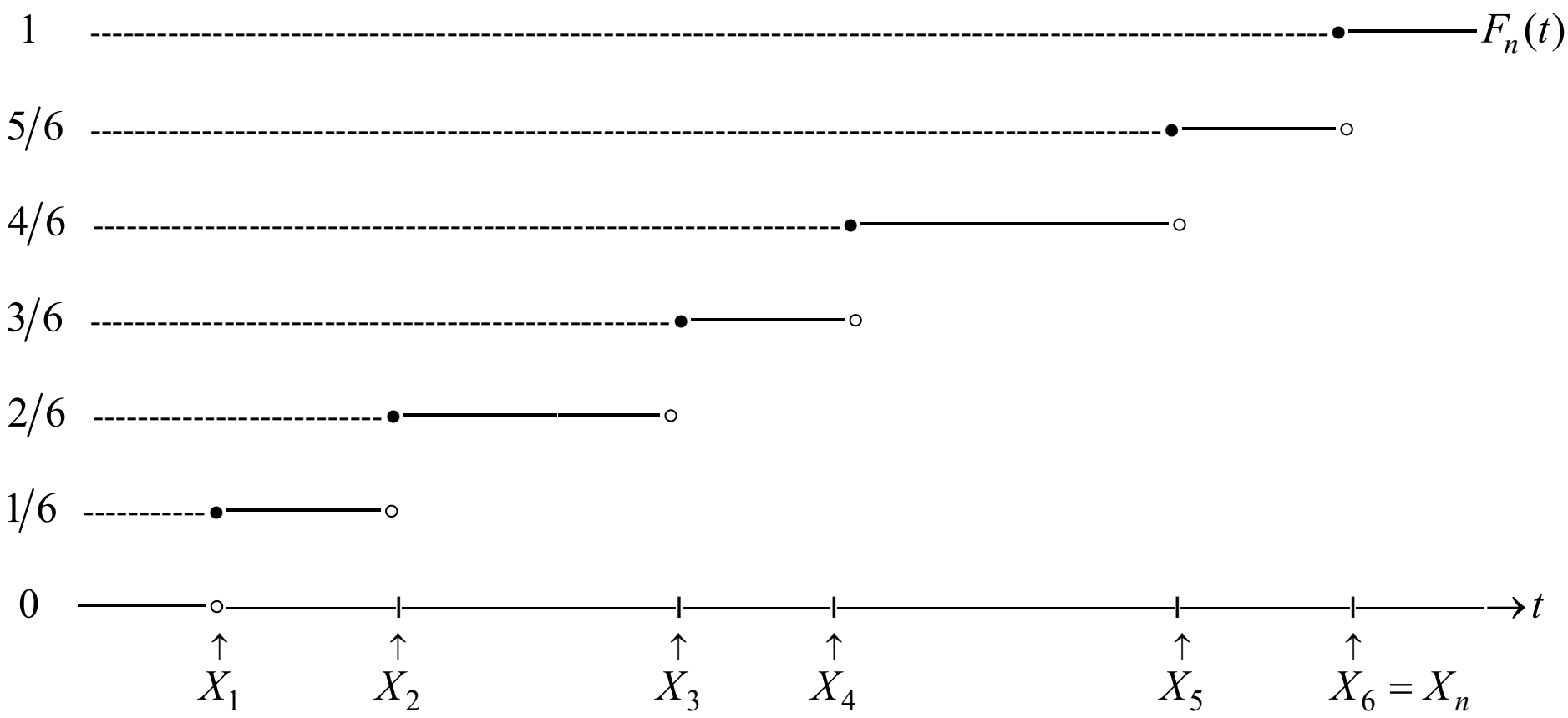
the empirical distribution functions (edf's) of $F(t)$ and $G(t)$ respectively. Then the *Two-sample Kolmogorov-Smirnov test* statistic is defined by

$$J = \frac{mn}{d}\sup_{t\in\mathbb{R}}\left|F_m(t) - G_n(t)\right| = \frac{mn}{d}\max_{1\le i\le N}\left|F_m(Z_{(i)}) - G_n(Z_{(i)})\right|, \qquad (4.24)$$

where $N = m + n$ and $Z_{(1)} \le \cdots \le Z_{(N)}$ are ordered values of $X_1,\ldots,X_m,Y_1,\ldots,Y_n$.

Empirical distribution function (edf) of a sample $X_1,\ldots,X_n$:

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le t\}}$$



The distribution of J

The exact distribution of J can be obtained in the following steps:

1. Arrange $X_1,\ldots,X_m$ and $Y_1,\ldots,Y_n$ in ascending order.
2. Take an XY sequence with m X's and n Y's, referred to as a *meshing*; e.g., $XYXYY$ is a meshing with $m=2$ and $n=3$.
3. Calculate $F_m(Z_{(1)}),\ldots,F_m(Z_{(N)})$ and $G_n(Z_{(1)}),\ldots,G_n(Z_{(N)})$ at $Z_{(i)}$ being an X or Y value according to the meshing taken in Step 2.
4. Find a value of J from $F_m(Z_{(1)}),\ldots,F_m(Z_{(N)}),G_n(Z_{(1)}),\ldots,G_n(Z_{(N)})$ and assign it probability $\binom{N}{n}^{-1}$.
5. Repeating Steps 1 – 4 for every meshing with given m and n produces the exact distribution of J under H_0:

$$\Pr(J=j)=\binom{N}{n}^{-1}(\text{No. of meshings that lead to } J=j)$$

Example 4.7 Let $m=2$, $n=3$, so that $d=\gcd(2,3)=1$. Consider $XYXYY$ and assume no ties. Then $F_m(t)=F_2(t)$ has jumps $1/2$ at $Z_{(1)}=X_1$ and $Z_{(3)}=X_2$; $G_n(t)=G_3(t)$ has jumps $1/3$ at $Z_{(2)}=Y_1$, $Z_{(4)}=Y_2$ and $Z_{(5)}=Y_3$. The values of $F_2(Z_{(i)})$ and $G_3(Z_{(i)})$ for $i=1,2,3,4,5$ are as follows:

$Z_{(i)}$	$Z_{(1)}=X_1$	$Z_{(2)}=Y_1$	$Z_{(3)}=X_2$	$Z_{(4)}=Y_2$	$Z_{(5)}=Y_3$
$F_2(Z_{(i)})$	$1/2$	$1/2$	1	1	1
$G_3(Z_{(i)})$	0	$1/3$	$1/3$	$2/3$	1

Then by (4.24),

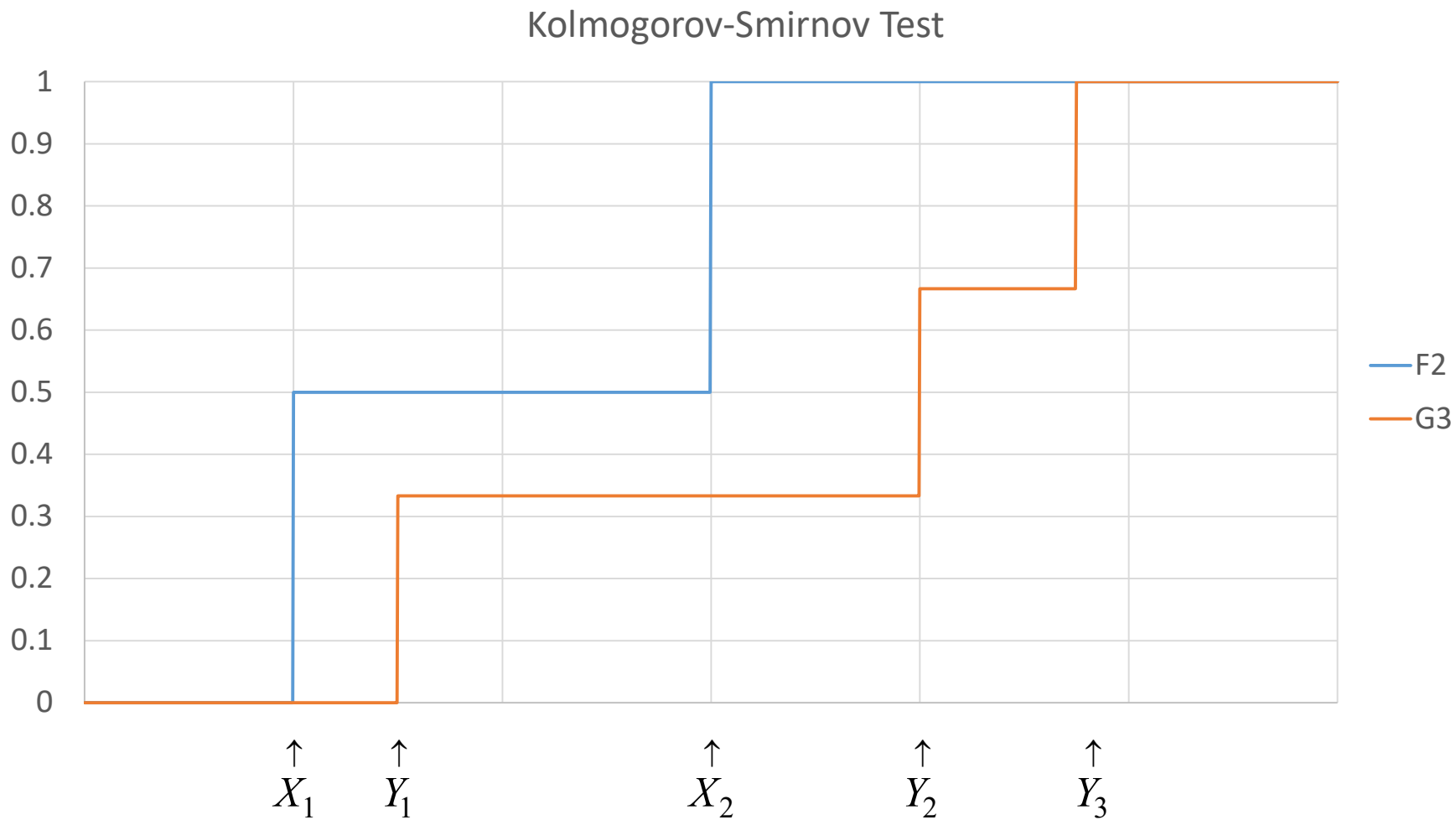
$$J=\frac{2\times 3}{1}\max_{1\le i\le 5}\left|F_2(Z_{(i)})-G_3(Z_{(i)})\right|=6\left|F_2(Z_{(3)})-G_3(Z_{(3)})\right|=6\left|1-\frac{1}{3}\right|=4$$

The probability of the J value from each meshing is $1/\binom{N}{n}=1/\binom{5}{3}=1/10=0.1$.

Repeat the above process for every meshing, we get the exact distribution of J:

j	2	3	4	6
$\Pr(J=j)$	0.1	0.3	0.4	0.2

Graphs of $F_2(t)$ and $G_3(t)$



Rejection rule: Reject H_0 at the (achievable) α level if $J \geq j_\alpha$, where

$$j_\alpha \text{ is a value of } J \text{ such that } \Pr(J \geq j_\alpha) = \alpha$$

It is determined by the exact distribution of J.

Asymptotic distribution of J: Under H_0, the limiting distribution of

$$J^* = \frac{d}{\sqrt{mnN}} J \tag{4.25}$$

is given by

$$\lim_{m,n\to\infty} \Pr(J^* \geq s) = \begin{cases} Q(s) & \text{if } s > 0 \\ 1 & \text{if } s \leq 0 \end{cases} \quad \text{with} \quad Q(s) = 1 - \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 s^2} \tag{4.26}$$

Approximate rejection rule: Reject H_0 at the α level if $J^* \geq q_\alpha^*$, where q_α^* is defined by $Q(q_\alpha^*) = \alpha$ and can be approximated by

$$q_\alpha^* \approx \sqrt{-0.5 \ln(\alpha/2)} \tag{4.27}$$

Example 4.8 The exact values of $q^*_{0.05}$ and $q^*_{0.01}$ by R are $q^*_{0.05} = 1.358$ and $q^*_{0.01} = 1.627$. Their approximations by (4.27) are quite close:

$$q^*_{0.05} \approx \sqrt{-0.5 \ln(0.05/2)} = \sqrt{1.8444} = 1.358$$

and

$$q^*_{0.01} \approx \sqrt{-0.5 \ln(0.01/2)} = \sqrt{2.6492} = 1.628$$

The p-value of the test based on the observed value j of J, or j^* of J^*, can be approximated by

$$p\text{-value} = \Pr(J \geq j) \approx \Pr(J^* \geq j^*) \approx 2e^{-2(j^*)^2} \qquad (4.28)$$

Ties: Ties will not affect the calculation of J. We can just follow the definition of the empirical distribution function (edf) to get the right answers. For example, if $Y_1 = Y_2 < Y_3$, then $G_3(Y_1) = G_3(Y_2) = 2/3$ instead of $G_3(Y_1) = 1/3$ for $Y_1 < Y_2 < Y_3$. The exact distribution of J will be affected by ties. But the rejection rules as above can still be applied approximately.

Example 4.9 Refer to Example 5.4 of the textbook (on page 192) with data:

$$(X_1,\ldots,X_{10}) = (-0.15, 8.60, 5.00, 3.71, 4.29, 7.74, 2.48, 3.25, -1.15, 8.38)$$

$$(Y_1,\ldots,Y_{10}) = (2.55, 12.07, 0.46, 0.35, 2.69, 0.94, 1.73, 0.73, -0.35, -0.37)$$

In this example, $m = n = 10$, $d = 10$, $N = 20$, and it is calculated that

$$\max_{1\le i\le 20} \left| F_{10}(Z_{(i)}) - G_{10}(Z_{(i)}) \right| = \frac{6}{10} = 0.6$$

Hence

$$J = \frac{mn}{d}(0.6) = \frac{10(10)}{10}(0.6) = 10(0.6) = 6$$

By (4.25) and (4.28),

$$J^* = \frac{10(6)}{\sqrt{10(10)(20)}} = \frac{3}{\sqrt{5}} \quad \text{and} \quad p\text{-value} = \Pr\left(J^* \ge \frac{3}{\sqrt{5}}\right) \approx 2e^{-2(9/5)} = 0.05465$$

The exact p-value by R is $\Pr(J \ge 6) = 0.05245$.