Review Article

BIOMEDICAL Journal of Scientific & Technical Research

ISSN: 2574 -1241

DOI: 10.26717/BJSTR.2021.35.005654

Moments and Cumulants of the Bivariate Mann-Whitney Statistic for Two-Stage Trials

Dewei Zhong and John Kolassa*

Department of Statistics and Biostatistics, the State University of New Jersey, USA

***Corresponding author:** John Kolassa, 565 Hill Center 110 Frelinghuysen Road Piscataway, NJ 08854, USA

ARTICLE INFO

Received: March 02, 2021

Published: April 16, 2021

Citation: Dewei Zhong, John Kolassa. Moments and Cumulants of the Bivariate Mann-Whitney Statistic for Two-Stage Trials. Biomed J Sci & Tech Res 35(1)-2021. BJSTR. MS.ID.005654.

ABSTRACT

This paper applies multivariate Cornish-Fisher techniques to calculate the asymptotic critical values of the bivariate Mann-Whitney statistic, which is used in two-stage study designs.

Introduction to Mann-Whitney Statistic and Two-Stage Test

Consider the problem of testing a difference in two groups. Suppose that the continuous responses $X_1, \ldots, X_M$ are from a control group, and the continuous responses $Y_1, \ldots, Y_N$ are from a treatment group. The Mann-Whitney U test [1], equivalent to the Wilcoxon rank sum test [2], uses the statistic

$$U = \sum_{j=1}^{N}\sum_{i=1}^{M} I_{ij}, \text{ for } I_{ij} = I(X_i < Y_j). \quad (1)$$

Here $I(X_i < Y_j)$ is 1 when $X_i < Y_j$ holds and 0 otherwise. The statistic U is designed to test the null hypothesis that the distribution of X_j is the same as that of Y_j, vs. the alternative hypothesis that $P[Y_j \geq X_j] > 0.5$, at level α. A critical value c is selected as the smallest value so that $P_0[U \geq c] \leq \alpha$. If U is larger than the critical value, the treatment group is determined to be superior to the control group.

Due to ethical concerns and resource management, common designs allow for early stopping in the presence of strong, early evidence. Spurrier and Hewett [3] provide a two-stage test based on the Mann-Whitney statistic. Wilding, et al. [4] discuss such a procedure in the context of clinical trials.

The two-stage test has two critical values, c_1 and c_2. First, gather m observations from control group and n observations from treatment group. Define

$$U_1 = \sum_{j=1}^{n}\sum_{i=1}^{m} I_{ij}. \quad (2)$$

If U_1 meets or exceeds the first critical value c_1, stop the trial early to declare the treatment group is superior to control group. If U_1 is less than c_1, gather $M-m$ observations from the control group and $N-n$ observations from treatment group, where $m \leq M$, $n \leq N$. Define

$$U_2 = \sum_{j=1}^{N}\sum_{i=1}^{M} I_{ij}. \quad (3)$$

If U_2 is larger than or equal the second critical value c_2, claim the treated is superior to the controls.

The critical value of Mann-Whitney statistic in one dimension can be easily calculated. The critical values for the two stage test are more difficult to calculate. Due to the complexity of the mass function for two dimensional Mann-Whitney statistics, obtaining exact critical values is computationally intensive. Kolassa, et al. [5] present a plan for approximating these critical values using a bivariate Cornish–Fisher expansion; this expansion requires bivariate cumulants of U_1 and U_2. Furthermore, they use a bivariate Edgeworth expansion to approximate power; this expansion also requires bivariate cumulants, in this case for an alternative distribution. This manuscript provides tools for calculating these bivariate moments, and hence bivariate cumulants. Under the null hypothesis, X_i and Y_i are jointly independent and identically distributed. The second section defines certain indicator functions and gives their null expectation. The third section presents first and second order joint moments of the Mann-Whitney statistics. The fourth section presents third- and fourth-order mixed moments. All of these moments are calculated under conditions general enough to encompass both the null and alternative distributions. The fifth section discusses the calculations of cumulants from moments.

Indicator Function Definitions and Expectations

Let I_{ij} take the value 1 if $X_i < Y_j$, and 0 otherwise. Products of these indicators represent indicators of more complicated sets. For example, $I_{ij}I_{il}I_{kj} = 1$ means that all of $X_i < Y_j$, $X_i < Y_l$, $X_k < Y_j$ hold, and $I_{ij}I_{il}I_{kj} = 0$ means that at least one of them does not hold. Below, moments of $U = (U_1, U_2)$ will be expressed as sums of such products. Terms will be factors with non-overlapping indices. Table 1 summarizes expectations of these factors. Zhong, et al. [6] perform these calculations in detail. Null values can be calculated using symmetry properties.

First- and Second-Order Moments

In general, using Table 1, $E[U_1] = \sum_{j=1}^{n}\sum_{i=1}^{m} E[I_{ij}] = mn\pi_0$, and $E[U_2] = MN\pi_0$.

Table 1: Expectations of Products of Indicators.

Probability	Definition	Null Value
π_0	$E[I_{ij}]$	1/2
π_1	$E[I_{ij}I_{kj}]$	1/3
π_2	$E[I_{ij}I_{kj}I_{aj}]$	1/4
π_3	$E[I_{ij}I_{kj}I_{aj}I_{ab}]$	3/20
π_4	$E[I_{ij}I_{kj}I_{ib}]$	5/24
π_5	$E[I_{ij}I_{kj}I_{ib}I_{ab}]$	2/15
π_6	$E[I_{ij}I_{kj}I_{aj}I_{cj}]$	1/5
π_7	$E[I_{ij}I_{kj}I_{ib}I_{id}]$	3/20
π_8	$E[I_{ij}I_{kj}I_{ib}I_{kb}]$	1/6
π_9	$E[I_{ij}I_{il}]$	1/3
π_{10}	$E[I_{ij}I_{kj}I_{id}I_{kb}]$	2/15
π_{11}	$E[I_{ij}I_{il}I_{ib}]$	1/4
π_{12}	$E[I_{ij}I_{il}I_{ib}I_{id}]$	1/5

Note that

$$U_2^2 = (\sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij})^2 = \sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij}^{\,2} + \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M} I_{ij}I_{kj} + \sum_{j=1}^{N}\sum_{i=1}^{M}\sum_{l=1,l\neq j}^{N} I_{ij}I_{il} + \sum_{j=1}^{N}\sum_{i=1}^{M}\sum_{l=1,l\neq j}^{N}\sum_{k=1,k\neq i}^{M} I_{ij}I_{kl}. \quad (4)$$

Substituting the probability from Table 1 into (4),

$$E[U_2^2] = MN\pi_0 + M(M-1)N\pi_1 + MN(N-1)\pi_9 + M(M-1)N(N-1)\pi_0^2. \quad (5)$$

 DOI: 10.26717/BJSTR.2021.35.005654

By the same reasoning,

$$E[U_1^2]=mn\pi_0+m(m-1)n\pi_1+mn(n-1)\pi_9+m(m-1)n(n-1)\pi_0^2. \tag{6}$$

Conditional expectations may be utilized to calculate mixed moments:

$$E[U_1U_2]=E[E[U_1U_2\,|\,U_2]]=E[U_2E[U_1U_2]]=E\left[U_2\frac{mn}{MN}U_2\right]=\frac{mn}{MN}E[U_2^2]. \tag{7}$$

Thus, by (6) and (7),

$$E[U_1U_2]=mn(\pi_0+(M-1)\pi_1+(N-1)\pi_9+(M-1)(N-1)\pi_0^2).$$

Higher Moments

As a tool for calculating $E[U_2^3]$ and $E[U_2^4]$ first define some sums that make up parts of this product. Let

$$C=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M}\sum_{l=1,l\neq j}^{N}\sum_{h=1,h\neq l,h\neq j}^{N} I_{ij}I_{il}I_{kh}U_2,$$

$$C^*=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M}\sum_{l=1,l\neq j}^{N}\sum_{g=1,g\neq i,g\neq k}^{M} I_{ij}I_{kj}I_{gl}U_2,$$

$$D=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M}\sum_{l=1,l\neq j}^{N}\sum_{h=1,h\neq l,h\neq j}^{N}\sum_{g=1,g\neq i,g\neq k}^{M} I_{ij}I_{kl}I_{gh}U_2,$$

$$E=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M}\sum_{l=1,l\neq j}^{N} I_{ij}I_{kj}I_{kl}U_2,$$

$$F=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M}\sum_{l=1,l\neq j}^{N} I_{ij}I_{kl}U_2,$$

$$G^*=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M} I_{ij}I_{kj}U_2,\quad G=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{l=1,l\neq j}^{N} I_{ij}I_{il}U_2,$$

$$H=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M} I_{ij}I_{kj}U_2^2,\quad H^*=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{l=1,l\neq j}^{N} I_{ij}I_{il}U_2^2,$$

$$K=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{l=1,l\neq j}^{N}\sum_{h=1,h\neq l,h\neq j}^{N} I_{ij}I_{il}I_{ih}U_2,$$

$$K^*=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M}\sum_{g=1,g\neq i,g\neq k}^{M} I_{ij}I_{kj}I_{gj}U_2,$$

$$L=\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M}\sum_{l=1,l\neq j}^{N} I_{ij}I_{kl}U_2^2=D+2C+2C^*+2F+2E.$$

Expectations of these sums of products of indicators can be calculated by separating the sums into quantities with indices replicated and independent quantities whose expectations are given in Table 1, to obtain:

$$E[C]=NM(M-1)(N-1)(N-2)[(M-2)(N-3)\pi_9\pi_0^2+3\pi_9\pi_0+(M-2)(2\pi_4\pi_0+\pi_9\pi_1)+(N-3)(\pi_{11}\pi_0+\pi_9^2)+\pi_7+2\pi_{10}],$$

$$E[C^*]=MN(M-1)(N-1)(M-2)[(M-3)(N-2)\pi_1\pi_0^2+3\pi_1\pi_0+(M-3)(\pi_2\pi_0+\pi_1^2)+(N-2)(2\pi_4\pi_0+\pi_1\pi_9)+\pi_3+2\pi_5],$$

$$E[D]=MN(M-1)(N-1)(M-2)(N-2)\pi_0(\pi_0^3(M-3)(N-3)+3(M-3)\pi_0\pi_1),$$

$$E[E]=NM(M-1)(N-1)[(M-2)(N-2)\pi_4\pi_0+(M-2)(\pi_3+\pi_5)+(N-2)(\pi_5+\pi_7)+3\pi_4+\pi_8],$$

$$E[F]=MN(M-1)(N-1)[\pi_0^3(M-2)(N-2)+2(M-2)\pi_0\pi_1+2(N-2)\pi_0\pi_9+2\pi_0^2+2\pi_4],$$

$$E[G]=MN(N-1)(\pi_9(M-1)(N-2)\pi_0+2(M-1)\pi_4+(N-2)\pi_{11}+2\pi_9),$$

$$E[G^*]=MN(M-1)(\pi_1(N-1)(M-2)\pi_0+2(N-1)\pi_4+(M-2)\pi_2+2\pi_1),$$

$$\begin{aligned}E[H]=MN(M-1)\{&\pi_1(M-2)(N-1)\times[\pi_0+(M-3)\pi_1+(N-2)\pi_9+(M-3)(N-2)\pi_0^2]+2(M-2)(M-3)(N-1)\pi_2\pi_0\\&+2(M-2)(N-1)\pi_3+4(N-1)(M-2)(N-2)\pi_4\pi_0+4(N-1)(M-2)\pi_5+4(M-2)(N-1)\pi_1\pi_0)+(M-2)(M-3)\pi_6\\&+2(N-1)(\pi_4+\pi_8)+2(N-1)(N-2)(\pi_7+\pi_{10})+4\pi_1+4(M-2)\pi_2+8(N-1)\pi_4+4(M-2)(N-1)\pi_3\},\end{aligned}$$

$$\begin{aligned}E[H^*]=MN(N-1)\{&\pi_9(M-1)(N-2)\times[\pi_0+(M-2)\pi_1+(N-3)\pi_9+(M-2)(N-3)\pi_0^2]+2(N-2)\pi_{11}(M-1)(N-3)\pi_0\\&+2(M-1)(N-2)\pi_7+4(M-1)(M-2)(N-2)\pi_4\pi_0+4(M-1)(N-2)\pi_{10}+4(M-1)(N-2)\pi_9\pi_0+(N-2)\pi_{11}\\&+(N-2)(N-3)\pi_{12}+2(M-1)(\pi_4+\pi_8)+2(M-1)(M-2)(\pi_3+\pi_5)+4\pi_9+4(N-2)\pi_{11}+8(M-1)\pi_4+4(N-2)(M-1)\pi_7\},\end{aligned}$$

$$E[K] = MN(N-1)(N-2)(\pi_{11}(M-1)(N-3)\pi_0 + 3(M-1)\pi_7 + (N-3)\pi_{12} + 3\pi_{11}),$$

$$E[K^*] = MN(M-1)(M-2)(\pi_2(N-1)(M-3)\pi_0 + (M-3)\pi_6 + 3(N-1)\pi_3 + 3\pi_2).$$

Then

$$U_2^3 = \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{M}\sum_{l=1}^{N} I_{ij} I_{kl} U_2$$

$$= \sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij}U_2 + \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M} I_{ij}I_{kj}U_2 + \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{l=1,l\neq j}^{N} I_{ij}I_{il}U_2 + \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M}\sum_{l=1,l\neq j}^{N} I_{ij}I_{kl}U_2$$

$$= U_2^2 + G^* + G + F,$$

and $E[U_2^3] = E[U_2^2] + E[G^*] + E[G] + E[F]$. Also,

$$U_2^4 = \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{M}\sum_{l=1}^{N} I_{ij} I_{kl} U_2^2$$

$$= \sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij}U_2^2 + \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M} I_{ij}I_{kj}U_2^2 + \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{l=1,l\neq j}^{N} I_{ij}I_{il}U_2^2 + \sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M}\sum_{l=1,l\neq j}^{N} I_{ij}I_{kl}U_2^2$$

$$= U_2^3 + H + H^* + L$$

$$= U_2^3 + H + H^* + D + 2C + 2C^* + 2F + 2E.$$

and $E[U_2^4] = E[U_2^3] + E[H] + E[H^*] + E[D] + 2E[C] + 2E[C^*] + 2E[F] + 2E[E]$.

Moments of U_1 are calculated substituting *m* and *n* for *M* and *N* respectively.

Conditional expectations are used to find mixed moments. In order to calculate expectations of mixed moments, introduce indicators indicating whether the observation ranked *i* in the first sample falls among those observations collected before the interim analysis, and similarly with the observation ranked *j* among the second sample:

$$A_i = \begin{cases} 1 & if\ X_{(i)} \in \{X_1, \ldots, X_m\} \\ 0 & otherwise \end{cases}$$

$$B_j = \begin{cases} 1 & if\ Y_{(j)} \in \{Y_1, \ldots, Y_{n_1}\} \\ 0 & otherwise \end{cases}$$

Then the Mann–Whitney statistic calculated using data before the interim analysis is

$$U_1 = \sum_{i=1}^{M}\sum_{j=1}^{N} A_i B_j I_{ij}.$$

The law of iterated expectations will be used to calculate mixed moments, by first conditioning on order statistics of the two samples ordered separately:

$$Z = (X_{(1)}, \ldots, X_{(M)}, Y_{(1)}, \ldots Y_{(N)}).$$

Then $E[A_i]$, $E[A_i, A_k]$ and $E[A_i, A_k, A_g]$, are $\lambda_x = \frac{m}{M}$, $\lambda_x^* = \frac{m(m-1)}{M(M-1)}$, $\lambda_x^\dagger = \frac{m(m-1)(m-2)}{M(M-1)(M-2)}$ respectively, with i, k, and g distinct integers among $\{1, \ldots m\}$. Similarly, $E[B_j]$, $E[B_j B_l]$, and $E[B_j B_l B_h]$, are $\lambda_y = \frac{n}{N}$, $\lambda_y^\dagger = \frac{n(n-1)(n-2)}{N(N-1)(N-2)}$, respectively, with j, l, and h distinct integers among $\{1, \ldots n\}$.

Calculation of mixed moments will proceed by expressing $U_1^T U_2^S$ in terms of quantities from U_1, *C*, *D*, *E*, *F*, *G*, *G**, *H*, *H**, *K*, and *K** as above, times

the indicators A_i, one such quantity attached to each distinct value of the first index, and times the indicators B_j, one such quantity attached to each distinct value of the second index. Then the expectations of products such as $A_i A_k$ with $i \neq k$ are expectations of products from a multinomial, and similarly with the B indicators. Then λ_x, λ_x^*, and $\lambda_x^\dagger$ are the expectations of products of one, two, and three such A, respectively, and λ_y, λ_y^*, and $\lambda_y^\dagger$ are the expectations of products of one, two, and three such *B*. Then

$$E[U_1 U_2] = E[E[U_1 U_2 \mid Z \mid] = E[U_2^2]\lambda_x \lambda_y.$$

Also,

 DOI: 10.26717/BJSTR.2021.35.005654

$$E[U_1^2U_2 \mid Z] = E\left[\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{M}\sum_{l=1}^{N} A_iB_jA_kB_l\, I_{ij}I_{kl}U_2 \mid Z\right]$$

$$= \lambda_x\lambda_y\sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij}U_2 + \lambda_x\lambda_x^*\lambda_y\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M} I_{ij}I_{kj}U_2$$

$$+\lambda_x\lambda_y\lambda_y^*\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{l=1,l\neq j}^{N} I_{ij}I_{il}U_2 + \lambda_x\lambda_x^*\lambda_y\lambda_y^*\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M}\sum_{l=1,l\neq j}^{N} I_{ij}I_{kl}U_2$$

$$= \lambda_x\lambda_y U_2^2 + \lambda_x\lambda_x^*\lambda_y G^* + \lambda_x\lambda_y\lambda_y^* G + \lambda_x\lambda_x^*\lambda_y\lambda_y^* F,$$

and

$$E[U_1^2U_2] = \lambda_x\lambda_y E[U_2^2] + \lambda_x\lambda_x^*\lambda_y E[G^*] + \lambda_x\lambda_y\lambda_y^* E[G] + \lambda_x\lambda_x^*\lambda_y\lambda_y^* E[F].$$

In the same way,

$$E[U_1^2U_2 \mid Z] = E\left[\sum_{i=1}^{M}\sum_{j=1}^{N} A_iB_jI_{ij}U_2^2 \mid Z\right] = \lambda_x\lambda_y U_2^3,$$

$$E[U_1U_2^2] = \lambda_x\lambda_y [U_2^3].$$

Next,

$$E[U_1^3U_2 \mid Z] = E\left[\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{M}\sum_{l=1}^{N}\sum_{g=1}^{M}\sum_{h=1}^{N} A_iB_jA_kB_lA_gB_hI_{ij}\, I_{kl}I_{gh}U_2 \mid Z\right]$$

$$= \lambda_x\lambda_y U_2^2 + 3\lambda_x\lambda_y^* G + \lambda_x\lambda_y^\dagger K + 3\lambda_x^*\lambda_y G^* + 3\lambda_x^*\lambda_y^* F$$

$$+6\lambda_x^*\lambda_y^* E + 3\lambda_x^*\lambda_y^\dagger C + \lambda_x^\dagger\lambda_y K^* + 3\lambda_x^\dagger\lambda_y^* C^* + \lambda_x^\dagger\lambda_y^\dagger D$$

and

$$E[U_1U_2^2] = \lambda_x\lambda_y E[U_2^2] + 3\lambda_x\lambda_y^* E[G] + \lambda_x\lambda_y^\dagger E[K] + 3\lambda_x^*\lambda_y E[G^*]$$

$$+3\lambda_x^*\lambda_y^* E[F] + 6\lambda_x^*\lambda_y^* E[E] + 3\lambda_x^*\lambda_y^\dagger E[C] + \lambda_x^\dagger\lambda_y E[K^*]$$

$$+3\lambda_x^\dagger\lambda_y^* E[C^*] + \lambda_x^\dagger\lambda_y^\dagger E[D].$$

Also

$$E[U_1^2U_2^2 \mid Z] = E\left[\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{M}\sum_{l=1}^{N} A_iB_jA_kB_l\, I_{ij}I_{kl}U_2^2 \mid Z\right]$$

$$= \lambda_x\lambda_y\sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij}U_2^2 + \lambda_x\lambda_y^*\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{l=1,l\neq j}^{N} I_{ij}I_{il}U_2^2$$

$$+\lambda_x^*\lambda_y\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M} I_{ij}I_{kj}U_2^2 + \lambda_x^*\lambda_y^*\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1,k\neq i}^{M}\sum_{l=1,l\neq j}^{N} I_{ij}I_{kl}U_2^2$$

$$= \lambda_x\lambda_y U_2^3 + \lambda_x\lambda_y^* H^* + \lambda_x^*\lambda_y H + \lambda_x^*\lambda_y^* L$$

$$= \lambda_x\lambda_y U_2^3 + \lambda_x\lambda_y^* H^* + \lambda_x^*\lambda_y H + \lambda_x^*\lambda_y^*(D + 2C + 2C^* + 2F + 2E),$$

$$E[U_1^2U_2^2] = \lambda_x\lambda_y E[U_2^3] + \lambda_x\lambda_y^* E[H^*] + \lambda_x^*\lambda_y E[H]$$

$$+\lambda_x^*\lambda_y^*(E[D] + 2E[C] + 2E[C^*] + 2E[F] + 2E[E])$$

Finally,

$$E[U_1U_2^3 \mid Z] = E\left[\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{M}\sum_{l=1}^{N} A_iB_j\, I_{ij}U_2^3 \mid Z\right]$$

$$= \lambda_x\lambda_y E\left[\sum_{i=1}^{M}\sum_{j=1}^{N} I_{ij}U_2^3 \mid Z\right] = \lambda_x\lambda_y E[U_2^4 \mid Z]$$

$$E[U_1U_2^3] = \lambda_x\lambda_y E[U_2^4].$$

Multivariate Cumulants

Multivariate cumulants can then be calculated from these moments. Let

$$\mu_{ij....k} = E[U_iU_j....U_k],$$

for indices i, j, ..., k taking values in {1, 2}. Define the moment generating function

$$M(T_1,T_2) = 1 + \sum_{l=1}^{\infty}\sum_{(i,j,....,k)\in\{1,2\}^l} \mu_{ij}...k^{T_iT_j...T_k/l!},$$

and define the cumulant generating function $K(T_1,T_2) = \log(M(T_1,T_2))$. Cumulants are then defined to be the coefficients $k_{ij....k}$ such that

$$M(T_1,T_2) = \sum_{l=1}^{\infty}\sum_{(i,j,....,k)\in\{1,2\}^l} k_{ij}...k^{T_iT_j...T_k/l!},$$

such that coefficients with indices permuted are equal. Analytic expressions for cumulants in terms of moments are simple in one dimension but are complex enough to be unusable in as few as two dimensions. Kolassa [7] presents software to perform these calculations numerically, as a result of using a symbolic calculus tool to output numerical code directly.

References

1. Mann HB, DR Whitney (1947) On a test of whether one of two random variables is stochastically larger than the other. Ann Math Statist 18(1): 50-60.
2. Wilcoxon F (1945) Individual comparisons by ranking methods. Biometrics Bulletin 1(6): 80-83.
3. Spurrier JD, JE Hewett (1976) Two-stage Wilcoxon tests of hypotheses. Journal of the American Statistical Association 71(356): 982-987.
4. Wilding GE, G Shan, AD Hutson (2012) Exact two-stage designs for phase ii activity trials with rank-based endpoints. Contemporary Clinical Trials 33(2): 332-341.
5. Kolassa J, X Chen, Y Seifu, D Zhong (2020) Power calculations and critical values for two-stage nonparametric testing regimes. Under review.
6. Zhong D, J Kolassa (2017) Moments and Cumulants of The Two-Stage Mann-Whitney Statistic. Technical report.
7. Kolassa J (2018) Two Stage: Two Stage MWW. R package version 1.0.

ISSN: 2574-1241

DOI: 10.26717/BJSTR.2021.35.005654

John Kolassa. Biomed J Sci & Tech Res