

Notes on Nonparametric Statistics

José Manuel Rodríguez Caballero

September 6, 2024

1 The Wilcoxon Rank Sum

Given two positive integers n and N , we define the Wilcoxon Rank Sum distribution $\mathbf{P}(W_n^N = k)$ via the generating function

$$\frac{\binom{N}{n}_q}{\binom{N}{n}} = \sum_{k=0}^{\frac{N(N+1)}{2}} \mathbf{P}(W_n^N = k) q^k, \quad (1)$$

where the polynomials $\binom{N}{n}_q$ are given by the product

$$\prod_{n=1}^N (1 - zq^n) = \sum_{n=0}^N (-1)^n \binom{N}{n}_q z^n. \quad (2)$$

Theorem 1. *For two positive integers n and N , we have*

$$\mathbf{E}(W_n^N) = \frac{n(N+1)}{2}, \quad (3)$$

$$\mathbf{Var}(W_n^N) = \frac{n(N-n)(N+1)}{12}. \quad (4)$$

Proof. Using the notations

$$F = \prod_{n=1}^N (1 - zq^n), \quad (5)$$

$$G = \sum_{k \geq 1} (q^k + q^{2k} + q^{3k} + \dots + q^{Nk}) \frac{z^k}{k}, \quad (6)$$

we have

$$F = e^{-G}. \quad (7)$$

Let $\mathcal{H} = q \frac{d}{dq}$, it follows

$$\mathcal{H}F = -e^{-G}\mathcal{H}, \quad (8)$$

$$\mathcal{H}^2 F = e^{-G} [(\mathcal{H}G)^2 - \mathcal{H}^2 G]. \quad (9)$$

We compute

$$\mathcal{H}G|_{q=1} = \frac{N(N+1)}{2} \frac{z}{1-z}, \quad (10)$$

$$\mathcal{H}^2 G|_{q=1} = \frac{N(N+1)(2N+1)}{6} \frac{z}{(1-z)^2}, \quad (11)$$

and

$$\mathcal{H}F|_{q=1} = -\frac{N(N+1)}{2} z(1-z)^{N-1}, \quad (12)$$

$$\mathcal{H}^2 F|_{q=1} = (1-z)^{N-2} \left[\left(\frac{N(N+1)}{2} \right)^2 z^2 - \frac{N(N+1)(2N+1)}{6} z \right]. \quad (13)$$

Considering the coefficients of both sides,

$$\frac{\mathcal{H} \binom{N}{n}_q|_{q=1}}{\binom{N}{n}} = \frac{n(N+1)}{2}, \quad (14)$$

$$\frac{\mathcal{H}^2 \binom{N}{n}_q|_{q=1}}{\binom{N}{n}} = \frac{n(N+1)(3nN+2n+N)}{12}. \quad (15)$$

By definition of the expected value,

$$\frac{\mathcal{H} \binom{N}{n}_q|_{q=1}}{\binom{N}{n}} = \mathbf{E} [W_n^N], \quad (16)$$

$$\frac{\mathcal{H}^2 \binom{N}{n}_q|_{q=1}}{\binom{N}{n}} = \mathbf{E} [(W_n^N)^2]. \quad (17)$$

Combining (14) and (16), we obtain (3).

By definition of the variance,

$$\mathbf{Var} [W_n^N] = \mathbf{E} \left[(W_n^N)^2 \right] - (\mathbf{E} [W_n^N])^2. \quad (18)$$

Combining (14), (15), (16) and (18) we obtain (4). \square