NOTES

An Upper Bound for the Expected Difference between Order Statistics

MANUEL LOPEZ
Rochester Institute of Technology
Rochester, NY 14623
malsma@rit.edu

JAMES MARENGO
Rochester Institute of Technology
Rochester, NY 14623
jemsma@rit.edu

Suppose we randomly and independently choose n numbers $X_1, X_2, \ldots, X_n$ from the interval $[0, 1]$ according to some probability distribution. Put these numbers in ascending order and call the results $Y_1 \le Y_2 \le \cdots \le Y_n$. If $1 \le k < \ell \le n$, how large can the number $Y_\ell - Y_k$ be? A moment's reflection reveals that by choosing the X's appropriately, we can make $Y_\ell - Y_k$ as small as zero or as big as one. But what if we consider the *expected value* of the random variable $Y_\ell - Y_k$? This expectation can be as small as zero if the common probability distribution of the X's is degenerate at a single point. But how large can this expectation be? We will answer that question in this article.

In statistics courses the set of random variables $X_1, X_2, \ldots, X_n$ is called a *random sample* and $Y_1, Y_2, \ldots, Y_n$ are called its *order statistics*. We represent the common probability distribution of X_i by the *cumulative distribution function* (cdf), defined for all real numbers x by $F(x) = P(X_i \le x)$. We assume that $F(0^-) \stackrel{\text{def}}{=} \lim_{x \to 0^-} F(x) = 0$ and $F(1) = 1$ so that $P(0 \le X_i \le 1) = 1$.

We first compute the probability distribution of Y_k, as follows. For $x \in [0, 1]$, let N_x be the number of observations among $X_1, X_2, \ldots, X_n$ which do not exceed x. The random variable N_x is therefore the number of successes in n Bernoulli trials, where each trial has success probability $p = F(x)$. Consequently, N_x has a *binomial distribution* and hence

$$\begin{aligned} P(Y_k > x) &= P(N_x \le k - 1) \\ &= \sum_{j=0}^{k-1} \binom{n}{j} [F(x)]^j \, [1 - F(x)]^{n-j} \, . \end{aligned} \tag{1}$$

The expected value of any random variable Y satisfying $P(0 \le Y \le 1) = 1$ is given by

$$\mathcal{E}(Y) = \int_0^1 P(Y > y) \, dy \tag{2}$$

(see, for example, Chung [**1**]). Note that this last integral always exists as a real number.

Math. Mag. **84** (2011) 365–369. doi:10.4169/math.mag.84.5.365.

It now follows from (1) and (2) that

$$\mathcal{E}_F\left(Y_\ell - Y_k\right) = \int_0^1 P_{k,\ell}\left(F(x)\right)\, dx, \tag{3}$$

where our notation indicates that this expectation depends on F and the polynomial $P_{k,\ell}$ is defined for $t \in [0, 1]$ by

$$P_{k,\ell}(t) = \sum_{j=k}^{\ell-1} \binom{n}{j} t^j \left(1-t\right)^{n-j}. \tag{4}$$

The uniform distribution

As an example, suppose that each X_i has the *Uniform distribution* with cdf given by $F(x) = x$ for $0 \le x \le 1$. It follows from (1) that

$$P(Y_k > x) = \sum_{j=0}^{k-1} \binom{n}{j} x^j (1-x)^{n-j}.$$

This distribution is called a *Beta distribution* (see for example, Hogg, McKean, and Craig [**2**]). We leave it as an exercise for the reader to show that

$$\int_0^1 x^j (1-x)^{n-j}\, dx = \frac{1}{(n+1)\binom{n}{j}}. \tag{5}$$

Readers who are familiar with Beta distributions will recognize this result. It now follows from (3), (4), and (5) that

$$\mathcal{E}\left(Y_\ell - Y_k\right) = \frac{\ell - k}{n+1}. \tag{6}$$

The reader should ask herself whether this last result seems intuitively reasonable.

We will next find a distribution for X_i (which will depend on the choice k and ℓ) that maximizes the value of $\mathcal{E}\left(Y_\ell - Y_k\right)$.

A Bernoulli distribution

Suppose each X_i has the probability distribution with point masses at zero (with probability p) and at one (with probability $1 - p$). This distribution is called a *Bernoulli distribution*. Its cdf depends on p and is given by

$$F_p(x) = \begin{cases} 0 & \text{if } x < 0 \\ p & \text{if } 0 \le x < 1 \\ 1 & \text{if } x \ge 1. \end{cases}$$

From (3) it then follows that

$$\mathcal{E}_p\left(Y_\ell - Y_k\right) = P_{k,\ell}(p).$$

Which value of p maximizes this expectation? After some algebra it is easy to see that

$$P'_{k,\ell}(p) = np^{k-1}(1-p)^{n-l}\left[\binom{n-1}{k-1}(1-p)^{\ell-k} - \binom{n-1}{\ell-1}p^{\ell-k}\right]. \tag{7}$$

Using the first derivative test it now follows $\mathcal{E}_p(Y_\ell - Y_k)$ is maximized for

$$p = p_{\max} \stackrel{\text{def}}{=} \left(1 + \left(\frac{\binom{n-1}{\ell-1}}{\binom{n-1}{k-1}}\right)^{\frac{1}{\ell-k}}\right)^{-1} \tag{8}$$

and that $\mathcal{E}_p(Y_\ell - Y_k) < \mathcal{E}_{p_{\max}}(Y_\ell - Y_k)$ if $p \neq p_{\max}$. For notational simplicity we have suppressed the dependence of $p_{\max}$ on k and ℓ.

The main result

The last example leads to our main result, which gives the maximum value for $\mathcal{E}_F(Y_\ell - Y_k)$ and answers the question posed in the introduction. We are maximizing this expectation over *all* possible probability distributions on [0, 1].

THEOREM 1. *Suppose $X_1, X_2, \ldots, X_n$ are independent random variables each having the same cdf F satisfying $F(0^-) = 0$ and $F(1) = 1$. For fixed integers k and ℓ satisfying $1 \le k < \ell \le n$, let Y_k and Y_ℓ be the kth and ℓth order statistics for $X_1, X_2, \ldots, X_n$. Define the polynomial $P_{k,l}$ as in (4), $p_{\max}$ as in (8) and $F_{\max}$ by*

$$F_{\max}(x) = \begin{cases} 0 & \text{if } x < 0 \\ p_{\max} & \text{if } 0 \le x < 1 \\ 1 & \text{if } x \ge 1. \end{cases}$$

Then $\mathcal{E}_F(Y_\ell - Y_k) \le P_{k,\ell}(p_{\max})$ with equality holding if and only if $F(x) = F_{\max}(x)$ for all x.

Proof. Using (3) and (7) and applying the first derivative test we have

$$P_{k,\ell}(p_{\max}) - \mathcal{E}_F(Y_\ell - Y_k) = \int_0^1 P_{k,\ell}(p_{\max}) - P_{k,\ell}(F(x))\,dx \ge 0. \tag{9}$$

Equality obviously holds if $F = F_{\max}$. Conversely, suppose there is a number $x_0 \in [0, 1)$ at which $F(x_0) \neq p_{\max}$. Since F is right continuous at x_0, it follows that there is a number $\delta > 0$ such that the integrand in (9) is strictly positive on $[x_0, x_0 + \delta)$. Since this integrand is nonnegative on [0, 1], the inequality in (9) must be strict in this case. ∎

A possible application and some examples

As a possible application of our theorem, suppose data are collected from some unknown distribution on the interval [0, 1] and the values of Y_k and Y_ℓ are obtained. Since $Y_\ell - Y_k$ is an unbiased estimator of its expectation, an observed value of this difference which grossly exceeds our upper bound may cast doubt on the assumption that our data are a random sample.

The upper bound $P_{k,\ell}(p_{\max})$ simplifies nicely in certain special cases. We explore two cases. The reader is invited to consider others.

CASE 1. $\ell = k + 1$

In this case $p_{\max} = \frac{k}{n}$ and so the maximum value for $\mathcal{E}_F(Y_{k+1} - Y_k)$ is

$$P_{k,k+1}(p_{\max}) = \binom{n}{k}\left(\frac{k}{n}\right)^k\left(1 - \frac{k}{n}\right)^{n-k}, \tag{10}$$

which is the binomial probability of obtaining exactly k successes in n Bernoulli trials each with success probability $\frac{k}{n}$.

To interpret this maximum value, we point out that according to our theorem, if the upper bound for $\mathcal{E}_F(Y_{k+1} - Y_k)$ is to be achieved, then every one of the X's must be either zero or one. If we interpret successes as zeros and failures as ones, then $Y_{k+1} - Y_k = 1$ when we have exactly k successes and $Y_{k+1} - Y_k = 0$ otherwise. Hence $\mathcal{E}_F(Y_{k+1} - Y_k)$ is the binomial probability of obtaining exactly k successes in n Bernoulli trials. It can easily be checked that the value of the success probability p which maximizes this binomial probability is $p_{\max} = \frac{k}{n}$.

CASE 2. $\ell = n + 1 - k$ where $k < \frac{n+1}{2}$

In this case $\binom{n-1}{k-1} = \binom{n-1}{\ell-1}$ so that $p_{\max} = \frac{1}{2}$ and hence the maximum value for the expected "trimmed range" $\mathcal{E}_F(Y_{n+1-k} - Y_k)$ is

$$\frac{1}{2^n}\sum_{j=k}^{n-k}\binom{n}{j} = 1 - \frac{1}{2^{n-1}}\sum_{j=0}^{k-1}\binom{n}{j}. \tag{11}$$

Our theorem shows that this maximum is achieved only in the case where

$$P(X_i = 0) = \frac{1}{2} = P(X_i = 1)$$

A different proof of that fact, which uses the notion of convexity, is given in [**3**] for the case $k = 1$.

Returning to the case where each X_i is uniformly distributed and recalling (6) we see from (10) that

$$\frac{1}{n+1} < \binom{n}{k}\left(\frac{k}{n}\right)^k\left(1 - \frac{k}{n}\right)^{n-k}$$

and from (11) that

$$\frac{k}{n+1} > \frac{1}{2^n}\sum_{j=0}^{k-1}\binom{n}{j}$$

We close by inviting the reader to verify these last two inequalities directly.

REFERENCES

1. K. L. Chung, *A Course in Probability Theory*, Harcourt, Brace & World, New York, 1968.
2. R. V. Hogg, J. W. McKean, and A. Craig, *Introduction to Mathematical Statistics*, 6th ed., Pearson Prentice Hall, Upper Saddle River NJ, 2005.
3. M. Lopez and J. Marengo, An upper bound for the expected range of a random sample, *College Math. J.* **41** (2010) 42–48.

Summary It is well known that the order statistics of a random sample from the uniform distribution on the interval [0, 1] have Beta distributions. In this paper we consider the order statistics of a random sample of n data points chosen from an arbitrary probability distribution on the interval [0, 1]. For integers k and ℓ with $1 \leq k < \ell \leq n$ we find an attainable upper bound for the expected difference between the order statistics Y_ℓ and Y_k. This upper bound depends on the choice of k and ℓ but does not depend on the distribution from which the data are obtained. We suggest a possible application of this result and we discuss some of its special cases.

Counting Irreducible Polynomials over Finite Fields Using the Inclusion-Exclusion Principle

SUNIL K. CHEBOLU
Illinois State University
Normal, IL 61790
schebol@ilstu.edu

JÁN MINÁČ
University of Western Ontario
London, ON N6A 5B7, Canada
minac@uwo.ca

Why there are exactly

$$\frac{1}{30}(2^{30} - 2^{15} - 2^{10} - 2^6 + 2^5 + 2^3 + 2^2 - 2)$$

irreducible monic polynomials of degree 30 over the field of two elements? In this note we will show how one can see the answer instantly using just very basic knowledge of finite fields and the well-known inclusion-exclusion principle.

To set the stage, let $\mathbb{F}_q$ denote the finite field of q elements. Then in general, the number of monic irreducible polynomials of degree n over the finite field $\mathbb{F}_q$ is given by Gauss's formula

$$\frac{1}{n}\sum_{d|n} \mu(n/d)q^d,$$

where d runs over the set of all positive divisors of n including 1 and n, and $\mu(r)$ is the Möbius function. (Recall that $\mu(1) = 1$ and $\mu(r)$ evaluated at a product of distinct primes is 1 or -1 according to whether the number of factors is even or odd. For all other natural numbers $\mu(r) = 0$.) This beautiful formula is well-known and was discovered by Gauss [**2**, p. 602–629] in the case when q is a prime.

We present a proof of this formula that uses only elementary facts about finite fields and the inclusion-exclusion principle. Our approach offers the reader a new insight into this formula because our proof gives a precise field theoretic meaning to each summand in the above formula. The classical proof [**3**, p. 84] which uses the Möbius' inversion formula does not offer this insight. Therefore we hope that students and users of finite fields may find our approach helpful. It is surprising that our simple argument is not available in textbooks, although it must be known to some specialists.

Math. Mag. **84** (2011) 369–371. doi:10.4169/math.mag.84.5.369.