DISTRIBUTION-FREE CONFIDENCE BOUNDS FOR $P(X<Y)$*

ZAKKULA GOVINDARAJULU

(Received Aug. 26, 1967)

1. Introduction and summary

Let X and Y denote two independent random variables having continuous F and G for their cumulative distribution functions, respectively. There are some practical situations (see, for instance, Birnbaum [1]) where one is interested in estimating $p=P(X<Y)$ on the basis of random samples $X_1, X_2, \cdots, X_m$ and $Y_1, Y_2, \cdots, Y_n$ from $F(x)$ and $G(y)$ respectively. It is well known that the minimum-variance unbiased estimate of p is the Mann-Whitney statistic given by $\hat{p}=\int_{-\infty}^{\infty} F_m(x)dG_n(x)$ where F_m and G_n denote the empirical distribution functions based on the random samples of sizes m and n of X and Y respectively. Then, the problem is, for given $\gamma(0<\gamma<1)$, to determine ε free of F and G such that either (i) $P(p\leqq\hat{p}+\varepsilon)=P(p\geqq\hat{p}-\varepsilon)\geqq\gamma$, or (ii) $P(|\hat{p}-p|\leqq\varepsilon)\geqq\gamma$. Birnbaum and McCarty [3] have obtained the inequality:

$$(1.1) \qquad \hat{p}-p\leqq\sup_x (F(x)-F_m(x))+\sup_x (G_n(x)-G(x))=D_m^-+D_n^+$$

where D_L^+ (D_L^-) is the Smirnov statistic based on a random sample of size L. Using the asymptotic distributions of $D_m^++D_n^+$, they obtained asymptotic values of $N^{1/2}\varepsilon$ for certain specified values of γ and m/N where $N=m+n$. Owen et al. [12] have extended Birnbaum and McCarty's [3] table of values of $N^{1/2}\varepsilon$ and observed that the same confidence bounds could conservatively be used when F and G are discontinuous. Also, for the problem of two-sided confidence bounds, one can easily obtain the inequality:

$$(1.2) \qquad |\hat{p}-p|\leqq\sup_x |F_m-F|+\sup_x |G_n-G|=D_m+D_n$$

where D_L is the Kolmogorov statistic based on a random sample of size L. However, using the (exact or asymptotic) distribution of D_m+D_n (See [7]), the present author finds the resultant bounds to be very crude. Hence, it is of interest to derive sharper confidence bounds for p.

* This research was supported by the National Science Foundation Grant NSF-GP 5664.

In this paper, the asymptotic normality of $\hat{p}$, when suitably standardized, is established when m and n tend to infinity. No restriction on the order of m and n is imposed, thus overcoming the criticism of Birnbaum [1] on the sufficient conditions of Lehmann [10] for the asymptotic normality of $\hat{p}$. A distribution-free upper bound for the asymptotic variance of $\hat{p}$ is derived which is equal to the bound of Van Dantzig [5]. One-sided and two-sided distribution-free confidence bounds for $\hat{p}-p$ based on the asymptotic normality and the distribution-free upper bound for the asymptotic variance are explicitly derived and they are approximately one half of the corresponding bounds due to Birnbaum and McCarty [3]. These distribution-free confidence bounds are found to be 80% as efficient as those based on normal samples of X and Y and obtained by Owen et al. [12] for the one-sided case and by Govindarajulu [8] for the two sided situation. An unbiased, consistent and distribution-free estimate of the asymptotic variance of $\hat{p}$ is derived which could be used instead of the upper bound, and thus shorten the confidence bounds.

2. Main results

The main results of this paper hinge upon the following theorem asserting the asymptotic normality of the Mann-Whitney statistic when either m or n tends to infinity. Asymptotic normality of $\hat{p}$, when suitably standardized, has been established by Mann and Whitney [11] for $F=G$. Lehmann [10] has shown that for all continuous F and G, $\hat{p}$ has an asymptotic normal distribution provided (i) $0<p<1$, and (ii) $n=cm$, $m\to\infty$. However, there will be some situations when m and n are not of the same order. Also, the asymptotic normality of $\hat{p}$ follows from the Chernoff-Savage theorem (see [4] or [9]) provided m/N is bounded away from zero and unity. Hence, the following theorem, although is a special case of the main theorem of [6], will be presented with a simple and independent proof.

THEOREM 2.1. *With the notation of section 1, for all continuous F and G, we have*

$$\lim_{m,n\to\infty} P\{\nu^{1/2}(\hat{p}-p)/\sigma\leqq z\}=\Phi(z) \tag{2.1}$$

where

$$\sigma^2=2\frac{\nu}{m}\iint_{x<y} F(x)[1-F(y)]dG(x)dG(y) + 2\frac{\nu}{n}\iint_{x<y} G(x)[1-G(y)]dF(x)dF(y) \tag{2.2}$$

$$\Phi(z)=(2\pi)^{-1/2}\int_{-\infty}^{z} e^{-t^2/2}dt\,, \tag{2.3}$$

and $\nu=\min(m, n)$ *provided* $\sigma^2\neq 0$.

PROOF. Write

$$\nu^{1/2}(\hat{p}-p)=\nu^{1/2}\int (F_m-F)dG+\nu^{1/2}\int Fd(G_n-G) + \nu^{1/2}\int (F_m-F)d(G_n-G)\,. \tag{2.4}$$

Now, integrating by parts once in the second term, we obtain $\nu^{1/2}(\hat{p}-p) = B_{m,n}+C_{m,n}$ where, after suppressing the subscripts,

$$B=\nu^{1/2}\left[\int (F_m-F)dG-\int (G_n-G)dF\right] \tag{2.5}$$

and

$$C=\nu^{1/2}\int (F_m-F)d(G_n-G)\,, \tag{2.6}$$

all the integrals being interpreted as ranging from $-\infty$ to ∞. From the central limit theorem we prove that B has an asymptotic normal distribution with mean zero and variance σ^2. Also, one can easily see that the mean of C is zero. Further one can rewrite C as

$$C=\frac{\nu^{1/2}}{n}\left[\sum_{j=1}^{n}\{F_m(Y_j)-F(Y_j)\}-\int (F_m-F)dG\right] \tag{2.7}$$

and obtain

$$E(C^2\mid X_1, X_2, \cdots, X_m)=\frac{\nu}{n}\int (F_m-F)^2dG-\frac{\nu}{n}\iint [F_m(x)-F(x)] \cdot [F_m(y)-F(y)]dG(x)dG(y)\,.$$

Thus

$$\begin{aligned}\operatorname{Var} C=EC^2&=\frac{\nu}{mn}\int F(1-F)dG\\ &\quad-\frac{2\nu}{mn}\iint_{x<y} F(x)[1-F(y)]dG(x)dG(y)\\ &\leqq \nu/4mn+\nu/4mn=\nu/2mn\,.\end{aligned} \tag{2.8}$$

Hence, C converges to zero in probability for all F and G as m and n tend to infinity. This completes the proof of the theorem.

Remark 2.1. Owing to its simplicity, proof of Theorem 2.1 is presented above, although Theorem 2.1 is contained in the theorem of Godwin and Zaremba [6]. Also, Theorem 2.1 follows from the proof of Hoeffding's theorem on U-statistics in the generalized version (for instance, see Lehmann [12] p. 964) with a change in the normalizing constant.

Remark 2.2. Suppose F and G are discontinuous having denumerable number of jump points (the jump points of F need not be different from those of G). If $F(G)$ has a jump of size α at a point t, remove the point t from the real line and insert in its place a closed interval of length α and distribute the probability mass α uniformly over this interval. The new distribution functions F^* and G^* so obtained are continuous. For samples obtained from F^* and G^* the relative order relations between X's and Y's have the same probability distribution as if the samples were drawn from F and G. Since $\hat{p}$ is well defined even if F and G have the same points of discontinuity, Theorem 2.1 is valid for all arbitrary F and G.

Remark 2.3. If m and n are random and there exists a positive integer N^* such that m/N^* and n/N^* converge to λ_1 and λ_2 in probability, then ν/N^* converges to $\lambda=\min(\lambda_1, \lambda_2)$ and $\mathcal{L}\{(\lambda N^*)^{1/2}(\hat{p}-p)/\sigma^* \leqq z\} \to \Phi(z)$ as $N^* \to \infty$ where

$$\sigma^{*2}=\frac{2\lambda}{\lambda_1}\iint_{x<y} F(x)[1-F(y)]dG(x)dG(y) + \frac{2\lambda}{\lambda_2}\iint_{x<y} G(x)[1-G(y)]dF(x)dF(y)\,. \tag{2.9}$$

Remark 2.4. Also, integrating by parts once in the first term on the right side of (2.4), we obtain, for the first order random term in $\nu^{1/2}(\hat{p}-p)$, $B=\nu^{1/2}\Big[\int Fd(G_n-G)-\int Gd(F_m-F)\Big]$. Consequently, the variance of B equals

$$\begin{aligned}\sigma^2&=\frac{\nu}{n}\Big[\int F^2dG-\Big\{\int FdG\Big\}^2\Big]+\frac{\nu}{m}\Big[\int G^2dF-\Big\{\int GdF\Big\}^2\Big]\\ &=\frac{\nu}{mn}\Big[m\Big\{\int F^2dG-p^2\Big\}+n\Big\{\int G^2dF-(1-p)^2\Big\}\Big]\,.\end{aligned} \tag{2.10}$$

Now, let $\nu=m$. Then,

$$\begin{aligned}n\sigma^2&=m\Big[\int F^2dG+\int G^2dF-p^2-(1-p)^2\Big]+(n-m)\Big[\int G^2dF-(1-p)^2\Big]\\ &=m\Big[\int (G-F)^2dF+\frac{2}{3}-p^2-(1-p)^2\Big]+(n-m)\Big[\int G^2dF-(1-p)^2\Big]\,.\end{aligned} \tag{2.11}$$

However, from Birnbaum and Klose [2] (see Lemma 3.1) we have the sharp inequality given by

$$\int (G-F)^2 dF \leqq \frac{1}{3} - p(1-p) . \tag{2.12}$$

Also, $\int G^2 dF - (1-p)^2 \leqq \int G dF - (1-p)^2 = p(1-p)$. Hence, it follows that

$$\sigma^2 \leqq \frac{1}{n} [mp(1-p) + (n-m)p(1-p)] = p(1-p) . \tag{2.13}$$

When $\nu = n$, one can analogously show that

$$\sigma^2 \leqq p(1-p) . \tag{2.14}$$

Thus, combining (2.13) and (2.14) we obtain

$$\sigma^2 \leqq p(1-p) \tag{2.15}$$

which is Van Dantzig's [4] bound for the variance of $\nu^{1/2}\hat{p}$. Notice that the expression for the exact variance of $\nu^{1/2}\hat{p}$ is different from σ^2 given by (2.10). Towards the distribution-free bound for σ^2 we obtain

$$\sigma^2 \leqq \frac{1}{4} . \tag{2.16}$$

Using Theorem 2.1 and the bound in (2.16) we readily obtain the following solutions for the distribution-free confidence bounds for p. For all F and G and non-random and large m or n, the solution ε of the equations

$$P(p \leqq \hat{p} + \varepsilon) = P(p \geqq \hat{p} - \varepsilon) \geqq \gamma , \qquad 0 < \gamma < 1 \tag{2.17}$$

is given by

$$\varepsilon \geqq (4\nu)^{-1/2} \Phi^{-1}(\gamma) \tag{2.18}$$

and the solution of the equation

$$P(|\hat{p} - p| \leqq \varepsilon) \geqq \gamma , \qquad 0 < \gamma < 1 \tag{2.19}$$

is given by

$$\varepsilon \geqq (4\nu)^{-1/2} \Phi^{-1}\left(\frac{1+\gamma}{2}\right) . \tag{2.20}$$

Remark 2.5. If m and n are random and there exists a positive integer N^* such that m/N^* and n/N^* converge respectively to λ_1 and λ_2, then the factor $(4\nu)^{-1/2}$ occuring in (2.18) and (2.20) should be replaced by $(4N^*\lambda)^{-1/2}$ where $\lambda=\min(\lambda_1, \lambda_2)$.

Remark 2.6. From eq. (2.5), it is clear that $\sigma^2\neq 0$ provided $0<p<1$. The case where either $p=0$ or $p=1$ is not of much practical interest.

Now, let us compare the one-sided confidence bounds given by (2.18) with those of Birnbaum and McCarty [3]. In Table 2.1, are presented values of $\delta=N^{1/2}\varepsilon$ for some specified values of γ when $m=n$. The values in parentheses are those of Birnbaum and McCarty [3].

Table 2.1 Showing values of $\delta=N^{1/2}\varepsilon$ for specified γ when $m=n=N/2$

γ	δ
.90	0.90 (2.65)
.95	1.17 (2.93)
.99	1.65 (3.49)
.995	1.82 (3.70)
.999	2.19 (4.14)

From Table 2.1, it is clear that the confidence bounds of Birnbaum and McCarty are at least twice as large as those given by eq. (2.18).

Also, it is of interest to compare the bounds given by (2.18) and (2.20) with those obtained by Owen et al. [12] and Govindarajulu [8] for the normal samples. These bounds for large $m=n$ are given by

$$\varepsilon\geqq 2\Phi\left(\frac{1}{2n^{1/2}}\Phi^{-1}(\gamma)\right)-1 \tag{2.21}$$

and

$$\varepsilon\geqq\Phi\left(n^{-1/2}\Phi^{-1}\left(\frac{1+\gamma}{2}\right)\right)-\frac{1}{2} \tag{2.22}$$

for the one-sided and two-sided situations respectively (see [12] and [8]). Also, when $m=n$, the bounds (2.18) and (2.20) respectively simplify to

$$\varepsilon=\Phi^{-1}(\gamma)/2n^{1/2}\,, \qquad \varepsilon=\Phi^{-1}\left(\frac{1+\gamma}{2}\right)\Big/2n^{1/2}\,. \tag{2.23}$$

For some specified values of n, and γ the bounds ε given in (2.23) are tabulated and are compared with those based on normal samples. The bounds based on normal equal-sized samples are given in parentheses.

Table 2.2 Showing values of ε for specified $n=m$ and γ

	γ \ $n=m$	25	64	100
One-sided	.95	.1645 (.1312)	.1028 (.0820)	.0833 (.0654)
	.99	.2330 (.1846)	.1456 (.1162)	.1167 (.0932)
Two-sided	.95	.1960 (.1524)	.1225 (.0968)	.0980 (.0777)
	.99	.2580 (.1950)	.1612 (.1265)	.1490 (.1016)

From Table 2.2, it is clear that the distribution-free bounds given by equations (2.18) and (2.20) are about 80% as efficient as those based on normal samples.

3. Distribution-free, unbiased, and consistent estimates of σ^2

One can obtain an unbiased, consistent and distribution-free estimate of σ^2 which can be used in the place of the upper bound*. Recall that $\frac{mn}{\nu}\sigma^2=m\left[\int F^2dG-p^2\right]+n\left[\int G^2dF-(1-p)^2\right]$, where $p=\int FdG$. Computations yield that

$$\text{(i)}\quad (m-1)^{-1}E\left[m\int F_m^2dG_n-\int F_mdG_n\right]=\int F^2dG\,,$$

$$\text{(ii)}\quad (n-1)^{-1}E\left[n\int G_n^2dF_m-\int G_ndF_m\right]=\int G^2dF\,,$$

$$\text{(iii)}\quad E\left\{\frac{mn}{(n-1)(m-1)}\left[\int F_mdG_n\right]^2-\frac{1}{(m-1)}\int(1-G_n)^2dF_m+\frac{1}{(m-1)(n-1)}\right.$$
$$\left.\cdot\int G_n^2dF_m-\frac{m}{(n-1)(m-1)}\int F_m^2dG_n-\frac{1}{(m-1)(n-1)}\int G_ndF_m\right\}=p^2\,,$$

(iv) $E\{$An expression obtained by interchanging the roles of F & G, and, m & n in (iii)$\}=(1-p)^2$.

The following computations will be helpful in verifying the above.

(a) $mF_m(x)=\chi_1(x)+\chi_2(x)+\cdots+\chi_m(x)$ where

$$\chi_i(x)=\begin{cases}1 & \text{when } x\geqq X \text{ (that is, with probability } F(x))\\ 0 & \text{when } x<X \text{ (that is, with probability } 1-F(x)),\end{cases}$$

* Sen [15] has considered distribution-free confidence limits for p by obtaining a similar distribution-free and consistent estimate of the variance of $\hat{p}$. The confidence bounds in section 3 are somewhat different from those of Sen [15].

with analogous statement when F & G, and m & n are interchanged.

(b) $E\left(\int F_m^2 dG_n\right)=\frac{1}{n}\sum EF_m^2(Y_j)=\frac{m-1}{m}\int F^2dG+\frac{1}{m}\int FdG$

(c) $$E\left[\int F_m dG_n\right]^2=\frac{m-1}{mn}\int F^2dG+\frac{2(n-1)}{mn}\int F(1-G)dG + \frac{(n-1)(m-1)}{mn}\left\{\int FdG\right\}^2$$

(d) $2\int F(1-G)dG=-\int Fd(1-G)^2=-\int(1-G)^2dF$

(e) $E\left[\int(1-G_n)^2dF_m\right]=\int(1-G)^2dF+n^{-1}\int GdF-n^{-1}\int G^2dF$.

Thus, an unbiased and distribution-free estimate of $\left\{\int F^2dG-p^2\right\}$ is given by

$$\frac{mn}{(m-1)(n-1)}\left[\int F_m^2dG_n-\left\{\int F_mdG_n\right\}^2\right]+\frac{1}{(m-1)}\left[\int(1-G_n)^2dF_m - \int F_mdG_n\right]+\frac{1}{(m-1)(n-1)}\left[\int G_ndF_m-\int G_n^2dF_m\right]. \tag{3.1}$$

Then, a consistent and distribution-free estimate of $\left\{\int F^2dG-p^2\right\}$ is given by $\left[\int F_m^2dG_n-\left\{\int F_mdG_n\right\}^2\right]$. By interchanging the roles of F & G, and m & n one can obtain, distribution-free, unbiased and consistent (or distribution-free and consistent) estimate of $\left\{\int G^2dF-(1-p)^2\right\}$. Thus a consistent estimate of σ^2 is given by $\hat{\sigma}^2$ where

$$\hat{\sigma}^2=\frac{\nu}{n}\left[\int F_m^2dG_n-\left\{\int F_mdG_n\right\}^2\right]+\frac{\nu}{m}\left[\int G_n^2dF_m-\left\{\int G_ndF_m\right\}^2\right] \tag{3.2}$$

using $\hat{\sigma}^2$ in the normal approximation, the confidence bounds for $p-\hat{p}$ are given by

$$\varepsilon\geqq(\hat{\sigma}^2/\nu)^{1/2}\Phi^{-1}(\gamma) \tag{3.3}$$

and

$$\varepsilon\geqq(\hat{\sigma}^2/\nu)^{1/2}\Phi^{-1}\left(\frac{1+\gamma}{2}\right) \tag{3.4}$$

for the one-sided and two-sided cases respectively. The bounds (3.3) and (3.4) would be shorter than (2.18) and (2.20) respectively.

4. Case when one of the distributions is known

There might be some practical situations where one of the distributions is exactly known and one is interested in setting up confidence bounds for $p=P(X<Y)$, on the basis of a random sample from the unknown population. Without loss of generality, assume that F is known and a random sample $Y_1, Y_2, \cdots, Y_n$ from G is available, where F and G are arbitrary cdf's. Then

$$\hat{p}=\int F dG_n . \tag{4.1}$$

One can easily obtain the following inequalities:

$$P(p<\hat{p}+\varepsilon)=P(p>\hat{p}-\varepsilon)\geqq P(D_n^+<\varepsilon)\geqq 1-e^{-2n\varepsilon^2}\geqq\gamma \tag{4.2}$$

and

$$P(|p-\hat{p}|<\varepsilon)\geqq P(D_n<\varepsilon)\geqq 1-2\sum_{j=1}^{\infty}(-1)^{j-1}e^{-2j^2\varepsilon^2 n}\geqq\gamma . \tag{4.3}$$

However, the solutions of ε for specified γ obtained from (4.2) and (4.3) would be cruder than the corresponding values of ε obtained from the following procedure. Applying the classical central limit theorem, we have

$$\lim_{n\to\infty} P[n^{1/2}(\hat{p}-p)/\sigma^*\leqq z]=\Phi(z) \tag{4.4}$$

where

$$\sigma^{*2}=\mathrm{Var}\, F(Y)=\int F^2(y)dG(y)-\left[\int FdG\right]^2 . \tag{4.5}$$

Further, $\sigma^{*2}\leqq p(1-p)\leqq 1/4$. Using this bound, one obtains (2.18) and (2.19) with $\nu=n$ as confidence bounds respectively for the one-sided and two-sided cases.

Also, one can easily see that an unbiased estimator of $\{EF(Y)\}^2$ is given by

$$n(n-1)^{-1}\left[\int F^2 dG_n-\left\{\int FdG_n\right\}^2\right] \tag{4.6}$$

and a consistent estimator of σ^{*2} is given by

$$\int F^2 dG_n-\left\{\int FdG_n\right\}^2 . \tag{4.7}$$

So, one obtains

$$\varepsilon\geqq\hat{\sigma}^* n^{-1/2}\Phi^{-1}(\gamma) , \qquad \varepsilon\geqq\hat{\sigma}^* n^{-1/2}\Phi^{-1}((1+\gamma)/2) \tag{4.8}$$

for the one- and two-sided confidence bounds respectively, where $\hat{\sigma}^*$ is given by (4.7).

5. Conclusion

Analogous distribution-free confidence bounds can be derived for $P(X<aY)$ and $P(X<h(Y))$ where a and h are given constant and function respectively and where X and Y could be vectors.

CASE INSTITUTE OF TECHNOLOGY*

REFERENCES

[1] Z. W. Birnbaum, "On a use of the Mann-Whitney statistics," *Proc. Third Berkeley Symposium on Mathematical Statistics and Probability*, 1, Univ. of Calif. Press, (1956), 13-17.

[2] Z. W. Birnbaum and Orval M. Klose, "Bounds for the variance of the Mann-Whitney statistic," *Ann. Math. Statist.*, 28 (1957), 933-945.

[3] Z. W. Birnbaum and R. C. McCarty, "A distribution-free upper confidence bound for $P(Y<X)$ based on independent samples of X and Y," *Ann. Math. Statist.*, 29 (1958), 558-562.

[4] Hermann Chernoff and I. Richard Savage, "Asymptotic normality and efficiency of certain nonparametric test statistics," *Ann. Math. Statist.*, 29 (1958), 972-994.

[5] D. Van Dantzig, "On the consistency and the power of Wilcoxon's two-sample test," *Nederl. Akad. Wetensch. Verslag Afd. Natuurk.*, Series A, 54, No. 1 (1951), 1-9.

[6] H. J. Godwin and S. K. Zaremba, "A central limit theorem for partly dependent variables," *Ann. Math. Statist.*, 32 (1961), 677-686 (especially, see pages 683-684).

[7] Zakkula Govindarajulu, "Asymptotic normality of a class of test statistics," unpublished manuscript, 1965.

[8] Zakkula Govindarajulu, "On the distribution of sum of two independent Kolmogorov statistics," *Sankhyā*, Series A, 29, Part I (1967), 61-82.

[9] Zakkula Govindarajulu, "Two-sided confidence limits for $P(X<Y)$ for normal samples of X and Y," to appear in *Sankhyā*, Series B, 29, Parts 1 & 2 (1967), 35-40.

[10] Z. Govindarajulu, L. Lecam and M. Raghavachari, "Generalizations of theorems of Chernoff-Savage on the asymptotic normality of test statistics," *Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1, Univ. of Calif. Press at Berkeley and Los Angeles, (1966), 609-638.

[11] E. L. Lehmann, "Consistency and unbiasedness of non-parametric tests," *Ann. Math. Statist.*, 22 (1951), 165-179.

[12] E. L. Lehmann, "Robust estimation in analysis of variance," *Ann. Math. Statist.*, 34 (1963), 957-966, (see § 6 appendix).

[13] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *Ann. Math. Statist.*, 18 (1947), 50-60.

[14] D. B. Owen, K. J. Craswell and D. L. Hanson, "Nonparametric upper confidence bounds for $P(Y<X)$ and confidence limits for $P(Y<X)$ when X and Y are normal," *J. Amer. Statist. Ass.*, 59 (1964), 906-924.

[15] P. K. Sen, "A note on asymptotically distribution-free confidence bounds for $P(X<Y)$ based on two independent samples," *Sankhyā*, Series A, 29, Part I (1967), 95-102.

* Now at University of Kentucky