AD-A083 529 CLEMSON UNIV S C DEPT OF MATHEMATICAL SCIENCES F/G 12/1
A NOTE ON THE INCOMPLETE BETA FUNCTION.(U)
MAY 79 K ALAM N00014-75-C-0451
UNCLASSIFIED N107 NL

1 OF 1
AD A
083529

END
DATE
FILMED
DTIC

12 LEVEL II

DEPARTMENT OF MATHEMATICAL SCIENCES

CLEMSON UNIVERSITY
Clemson, South Carolina

80 4 24 002

A NOTE ON THE
INCOMPLETE BETA FUNCTION

By

Khursheed Alam

Clemson University

Report N107

Technical Report #316

May 25, 1979

Research Supported in part by
THE OFFICE OF NAVAL RESEARCH

Task NR 047-202 Contract N00014-75-C-0451

A Note On The Incomplete Beta Function

Khursheed Alam*

Clemson University

Abstract

The incomplete beta function arises in various statistical problems. It is known, for example, that the tail probability of the binomial distribution can be expressed as an incomplete beta function. This paper gives some results on a monotonicity property of the incomplete beta function. The given results are shown to have application in a problem of ranking and selection.

Key words: Binomial Distribution; Ranking & Selection

AMS Classification: 62E99

*The author's work was supported by the Office of Naval Research under Contract N00014-75-C-0451

1. Main results. There are given below two theorems on the monotonicity property of an incomplete beta function. An application of the given result is shown in the next section. Let

$$I_p(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^p x^{a-1} (1-x)^{b-1} dx,$$

$$a > 0,\ b > 0,\ 0 < p < 1$$

denote the incomplete beta function. Let $a = \xi n+\gamma$ and $b = (1-\xi)n+\delta$, where $0 < \xi > 1$, $n > 0$, $\gamma \geq 0$, $\delta \geq 0$. The following theorem establishes a monotonicity property of $I_p(\xi n+\gamma,\ (1-\xi)n+\delta)$ in terms of n. Let $n' > n$ and

$$f(p) = I_p(n'\xi+\gamma,\ n'(1-\xi)+\delta) - I_p(n\xi+\gamma,\ n(1-\xi)+\delta).$$

Theorem 1.1. Let $0 < p_o < 1$. If $f(p_o) \leq (\geq)\ 0$ then $f(p) \leq (\geq) 0$ for $p \leq (\geq) p_o$.

Proof: Let

$$g_n(x) = x^{n\xi+\gamma-1} (1-x)^{n(1-\xi)+\delta-1} / B(n\xi+\gamma,\ n(1-\xi)+\delta)$$

denote the beta density function, and let

$$h(x) = g_{n'}(x) / g_n(x).$$

Clearly, $h(x)$ is nondecreasing (nonincreasing) in x for $x \leq \xi (\geq) \xi$.

Consider the function

$$f(p) = \int_0^p (h(x)-1) g_n(x)\, dx. \tag{1.1}$$

As x varies from 0 to 1, the integrand on the right hand side of (1.1) is either negative throughout or it changes sign from negative to positive and then to negative. Since $f(1) = 0$, it follows that $f(p)$ changes sign once from negative to positive as p varies from 0 to 1. The conclusion of the theorem follows immediately. []

Let $\gamma = 0$, $\delta = 1$. If n and $n\xi$ are integer valued then

$$(1.2) \qquad I_p(n\xi,\ n(1-\xi)+1) = \sum_{r=\xi n}^{n} \binom{n}{r} p^r (1-p)^{n-r}$$

represents the tail probability of a binomial distribution. Let $\psi(n) = \partial \log \Gamma(n)/\partial n$ denote the digamma function, and let

$$A(n) = \psi(n) - \xi\psi(n\xi) - (1-\xi)\psi(n(1-\xi)) + \xi\log\xi + (1-\xi)\log(1-\xi) - 1/n.$$

Using the integral formula for the digamma function, given by

$$\psi(n) = \log n - \frac{1}{2n} - 2\int_0^\infty t(e^{2\pi t}-1)^{-1}(t^2+n^2)^{-1}dt$$

we have

$$(1.3) \qquad A(n) = 2\int_0^\infty t(e^{2\pi t}-1)^{-1}\left(\frac{\xi}{t^2+n^2\xi^2} + \frac{1-\xi}{t^2+n^2(1-\xi)^2} - \frac{1}{t^2+n^2}\right)dt - \frac{1}{2n}$$

$$\leq \pi^{-1}\int_0^\infty\left(\frac{\xi}{t^2+n^2\xi^2} + \frac{1-\xi}{t^2+n^2(1-\xi)^2} - \frac{1}{t^2+n^2}\right)dt - \frac{1}{2n}$$

$$= 0\ .$$

Consider the function

$$C(x) = \xi \log x + (1-\xi)\log(1-x) - \xi\log\xi - (1-\xi)\log(1-\xi) + \frac{1}{n}$$

$$0 < x < 1.$$

Clearly $C(x)$ is a concave function of x. Let p_n and q_n denote the roots of the equation $C(x) = 0$, where $p_n < \xi < q_n$. Note that p_n and $q_n \to \xi$ as $n \to \infty$. We have

$$(1.4) \qquad \partial^2 I_p(\xi n,\ (1-\xi)n+1)/\partial p \partial n = \frac{p^{n\xi-1}(1-p)^{n(1-\xi)}}{B(n\xi, n(1-\xi)+1)} [\xi \log p + (1-\xi)\log(1-p) + \psi(n+1) - \xi\psi(n\xi) - (1-\xi)\psi(n(1-\xi)+1)]$$

$$= \frac{p^{n\xi-1}(1-p)^{n(1-\xi)}}{B(n\xi, n(1-\xi)+1)} [C(p) + A(n)].$$

The second equality in (1.4) follows from the relation $\psi(n+1) = \psi(n) + \frac{1}{n}$. In view of (1.3) we have that the right hand side of (1.4) is negative for $p \le p_n$ and $p \ge q_n$.

Since $\partial I_p(n\xi, n(1-\xi)+1)/\partial n \to 0$ as $p \to 0$, it follows that

$$(1.5) \qquad \partial I_p(n\xi, n(1-\xi)+1)/\partial n < 0$$

for $p \le p_n$. Using the relation $I_p(n\xi, n(1-\xi)+1) = 1 - I_{1-p}(n(1-\xi)+1, n\xi)$, we find that the reverse inequality holds in (1.5) for $p \ge q_n$. We have proved the following result.

<u>Theorem 1.2</u>. The incomplete beta function $I_p(n\xi, n(1-\xi)+1)$ is

decreasing in n for $p \le p_n$ and increasing in n for $p \ge q_n$, where p_n and q_n are the roots of the equation $C(x) = 0$.

Remark 1.1. It can be shown that the monotonicity property given by Theorem 1.2,holds also for the function $I_p(n\xi, n(1-\xi))$, where

$$I_p(n\xi, n(1-\xi)) = \sum_{r = \xi n}^{n} \binom{n-1}{r} p^r (1-p)^{n-1-r}$$

if n and $n\xi$ are positive integers.

Remark 1.2. From Formula (1e.6.2) of Rao (1966) we have

$$C(x) \le \frac{1}{n} - \frac{(x-\xi)^2}{2} .$$

Therefore, $0 < \xi - p_n \le \sqrt{\frac{2}{n}}$ and $0 < q_n - \xi \le \sqrt{\frac{2}{n}}$.

2. Application. Consider the following problem of ranking and selection. There are given k populations with cumulative distribution functions (cdf) $F_i(x) = F_i$ $(i=1,\ldots,k)$, and a number α with $0 < \alpha < 1$. The distribution functions are unknown but they are assumed to be continuous. Let ξ_i^α denote the α-quantile of F_i. It is assumed for simplicity that ξ_i^α is uniquely determined for each $i = 1,\ldots,k$. Given a sample of n observations from each population, it is required to select the population associated with the largest value of ξ_i^α, called the "best" population. We snall assume that $n\alpha$ is integer valued.

Let x_{ij} denote the jth order statistic in the sample from F_i, and let $j = n\alpha$. Suppose that the population associated with the largest value of x_{ij} is selected as the best population. Let the ith population be the best population. Then the probability of a correct selection (PCS) is given by

$$(2.1) \qquad PCS = n\binom{n-1}{j-1} \int_0^1 u^{j-1}(1-u)^{n-j} \prod_{\substack{t=1 \\ t \neq i}}^{k} I_{F_t F_i^{-1}(u)}(n\alpha, n(1-\alpha)+1)\,du.$$

For large n, the right hand side of (2.1) is approximately given by

$$(2.2) \qquad \prod_{\substack{t=1 \\ t \neq i}}^{k} I_{F_t(\xi_i^{\alpha})}(n\alpha, n(1-\alpha)+1).$$

If it is assumed that the α-quantile of the best population is sufficiently larger than the α-quantile of each of the remaining populations, in the sense that

$$F_t\,(\xi_i^{\alpha}) \geq \alpha+\varepsilon \;, \; t \neq i$$

where ε is a given positive number, then (2.2) is minimized for

$$F_t(\xi_i^{\alpha}) = \alpha + \varepsilon, \quad t \neq i.$$

Therefore, the minimum probability of a correct selection is approximately given by

$$(2.3) \qquad \min PCS = (I_{\alpha+\varepsilon}(n\alpha, n(1-\alpha)+1))^{k-1}$$

The right hand side of (2.2) is increasing in n for $\varepsilon > \frac{2}{\sqrt{n}}$ by Theorem 1.2 and Remark 1.2. Thus a minimum value of n can be determined for the given selection problem, for which the probability of a correct selection is at least as large as a given number P^* $(\frac{1}{k} < P^* < 1)$.

The problem of selecting the best population for the largest α-quantile, has been considered by Rizvi and Sobel (1967) and Sobel (1967).

In the application given above, the problem of selecting the population associated with largest median value, that is, when $\alpha = \frac{1}{2}$, is of special interest. For this case the quantity inside the square bracket on the right hand side of (1.4) is given by

$$C(p) + A(n) = \psi(n) - \psi(\tfrac{n}{2}) + \tfrac{1}{2}\log(p(1-p)) \tag{2.4}$$

$$= \tfrac{1}{2}(\psi(\tfrac{n+1}{2}) - \psi(n) + \log(4p(1-p))).$$

Let p_o and $1-p_o$ be the values of p obtained by equating the right hand side of (2.4) to zero. Then $I_p(\frac{n}{2}, \frac{n}{2}+1)$ is decreasing in n for $p < p_o$ and increasing in n for $p > 1-p_o$. To illustrate our result, we give below values of $I_p(\frac{n}{2}, \frac{n}{2}+1)$ for $p = (.5, .55)$ and

$n = 2, 4, 8, (2)14, 20, 100$. It appears from the table that $I_{1/2}(\frac{n}{2}, \frac{n}{2}+1)$ decreases as n varies from 2 to 10 and increases thereafter.

$$I_p(\tfrac{n}{2}, \tfrac{n}{2}+1)$$

n =	2	4	8	10	12	14	20	100
p = .50	.7500	.6875	.6367	.6230	.6128	.6047	.5881	.5398
p = .55	.7975	.7585	.7396	.7384	.7393	.7414	.7505	.8654

References

[1] Rao, C. R. (1966). Linear Statistical Inference. Wiley Publications in Statistics.

[2] Rizvi, M. H. and Sobel, M. (1967). Nonparametric procedures for selecting a subset containing the population with the largest α-quantile. Ann. Math. Statist. (38) 1788-1303.

[3] Sobel, M. (1967). Nonparametric procedures for selecting the t populations with the largest α-quantiles. Ann. Math. Statist. (38) 1804-1816.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER N-107	2. GOVT ACCESSION NO. AD-A083529	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) "A Note On The Incomplete Beta Function"		5. TYPE OF REPORT & PERIOD COVERED
		6. PERFORMING ORG. REPORT NUMBER Technical Report #316
7. AUTHOR(s) Khursheed Alam		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0451
9. PERFORMING ORGANIZATION NAME AND ADDRESS Clemson University, Dept. of Mathematical Sciences, Clemson, South Carolina 29631		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 047-202
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research, Code ~~436~~ 434, Arlington, Va. 22217		12. REPORT DATE May 25, 1979
		13. NUMBER OF PAGES 9
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release: distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Binomial Distribution; Ranking & Selection

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The incomplete beta function arises in various statistical problems. It is known, for example, that the tail probability of the binomial distribution can be expressed as an incomplete beta function. This paper gives some results on a monotonicity property of the incomplete beta function. The given results are shown to have application in a problem of ranking and selection.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)