(TRADITIONAL) EXACT CONFIDENCE INTERVALS FOR THE BINOMIAL DISTRIBUTION

L. THOMAS RAMSEY
1521 ALEXANDER ST APT 503
HONOLULU, HI 96822

Abstract. We describe the traditional form of exact confidence intervals for the binomial distribution.

1. Introduction

Consider $X(n,p)$, a binomial random variable with positive integer parameter n and success parameter $p \in [0,1]$. For integers $k \in [0,n]$,

$$P(X(n,p)=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

For all other real numbers r, $P(X(n,k)=r)=0$. Hereafter $[0\ldots n]$ denotes the set of integers in the real interval $[0,n]$, with similar meanings assigned to $[a\ldots b)$, $(a\ldots b)$, etc. Throughout, n denotes a positive integer.

2. Probabilities of "Tail" Intervals

For $j \in [0\ldots n]$, let the lower binomial tail be this function:

$$h(j,p) = P(X \le j) = \sum_{k=0}^{j} \binom{n}{k} p^k (1-p)^{n-k}.$$

Note that, for $j=n$ and all $p \in [0,1]$,

$$h(n,p) = [p + (1-p)]^n = 1.$$

The next proposition says that, except when $j=n$, (nontrivial) lower tails have probabilities that strictly decrease with p.

Proposition 1. *Let $j \in [0\ldots n)$.*

- *$h(j,0)=1$ and $h(j,1)=0$.*
- *For $0<p<1$, one has*

$$\frac{\partial h(j,p)}{\partial p} = -n\binom{n-1}{j} p^j (1-p)^{(n-1)-j} < 0.$$

- *$h(j,p)$ is strictly decreasing for $p \in [0,1]$.*

Date: December 17, 2003.

Proof: When $p = 0$, $h(j, 0)$ reduces to its first term $(1 - p)^n = 1$. Because $j < n$, all terms of $h(j, p)$ have a positive power of $1 - p$. Since $1 - p = 0$ when $p = 1$, we have $h(j, 1) = 0$.

Let $h'(j, p)$ denote the partial derivative of $h(j, p)$ with respect to p.

When $j = 0$ we have $h(j, p) = (1 - p)^n$ and hence

$$h'(j,p) = -n(1-p)^{n-1} = -n\binom{n-1}{0}p^0(1-p)^{n-1-0} < 0 \text{ for } p \in (0,1)$$

Since h is continuous in p, we have h strictly decreasing on $[0, 1]$.

Now suppose that $j > 0$ (and hence $n > 1$). Then

$$\begin{aligned} h'(j,p) &= n(1-p)^{n-1}(-1) \\ &\quad + \textstyle\sum_{k=1}^{j} \binom{n}{k}[kp^{k-1}(1-p)^{n-k} + p^k(n-k)(1-p)^{n-k-1}(-1)] \\ &= -n(1-p)^{n-1} + n\left[\textstyle\sum_{k=1}^{j}\binom{n-1}{k-1}p^{k-1}(1-p)^{n-1-(k-1)}\right] \\ &\quad -n\left[\textstyle\sum_{k=1}^{j}\binom{n-1}{k}p^k(1-p)^{(n-1)-k}\right] \\ &= -n(1-p)^{n-1} + n\left[\textstyle\sum_{k=0}^{j-1}\binom{n-1}{k}p^k(1-p)^{(n-1)-k}\right] \\ &\quad -n\left[\textstyle\sum_{k=1}^{j}\binom{n-1}{k}p^k(1-p)^{(n-1)-k}\right] \\ &= -n(1-p)^{n-1} + n(1-p)^{n-1} - n\binom{n-1}{j}p^j(1-p)^{(n-1)-j} \\ &= -n\binom{n-1}{j}p^j(1-p)^{(n-1)-j}. \end{aligned}$$

For $p \in (0, 1)$, $h'(j, p) < 0$. Since $h(j, p)$ is continuous in p, we have $h(j, p)$ strictly decreasing on $[0, 1]$. □

Similarly, the upper binomial tail is defined to be this function:

$$w(j,p) = P(X \geq j) = \sum_{k=j}^{n}\binom{n}{k}p^k(1-p)^{n-k}.$$

For $j > 0$, we have $w(j, p) = 1 - h(j - 1, p)$. For $j = 0$ and for all $p \in [0, 1]$,

$$w(0,p) = [p + (1-p)]^n = 1.$$

The next proposition says that, except when $j = 0$, (nontrivial) upper tails have probabilities that strictly increase with p.

Proposition 2. *Let* $j \in [1 \dots n]$.

- $w(j, 0) = 0$ *and* $w(j, 1) = 1$.
- *For* $0 < p < 1$, *one has*
$$\frac{\partial w(j,p)}{\partial p} = n\binom{n-1}{j-1}p^{j-1}(1-p)^{n-j} > 0.$$
- $w(j, p)$ *is strictly increasing for* $p \in [0, 1]$.

Proof: With $j > 0$, we have $w(j, p) = 1 - h(j - 1, p)$. When $j \in [1 \dots n]$, we have $j - 1 \in [0 \dots n - 1]$. Application of the previous proposition to $h(j - 1, p)$ gives the conclusions listed here. □

3. Traditional Exact Confidence Intervals

Let $p \in [0, 1]$ and n a positive integer. Given a number $\alpha \in (0, 1)$ and an "observation" $k \in [0 \dots n]$ from a binomial random variable $X(n, p)$, we would like to describe a $(100 \times \alpha)\%$ confidence interval $I(n, k) \subset [0, 1]$. The basic idea behind traditional exact confidence intervals is this: an observation like k should be an unusual observation for any p outside of $I(n, k)$. Traditionally, an observation like k translates into an observation in $[0, k]$ or $[k, n]$. "Unusual" translates into $P(X(n,p) \in [0,k]) \leq (1-\alpha)/2$ or $P(X(n,p) \in [k,n]) \leq (1-\alpha)/2$. Since $\alpha \in (0,1)$, both cannot be true, and one might logically replace $(1-\alpha)/2$ with $1-\alpha$. The tradition is otherwise. Here is a formal statement of the traditional translation of the basic idea:

$$p \notin I(n,k) \Rightarrow [P(X(n,p) \leq k) \leq (1-\alpha)/2 \quad \text{or} \quad P(X(n,p) \geq k) \leq (1-\alpha)/2]$$

Another way to state the idea is that, having observed k, any $p \notin I(n,k)$ is "rejected" but with a risk: for any specific rejected p but which is in fact correct, the probability of obtaining "something like" k from $X(n,p)$ is at most $1-\alpha$. Thus, with M repeated, independent observations from $X(n,p)$, with M large, by the central limit theorem we would expect to reject p about $(1-\alpha)M$ times and make the correct decision to not reject p about αM times. Repeated bad luck (getting "something like" k when it is unusual for the true value of p) is unlikely. By choosing α we can restrict the expected frequency for mistaken decisions.

One unenlightening strategy would be to set $I(n,k) = [0,1]$. No p would be rejected and one would never wrong (p is indeed always in $[0,1]$). However, by accepting a certain risk level, and reporting a "confidence interval" with a specific α and thus announcing that we could be wrong with an expected frequency bounded by $1-\alpha$, we might as well make $I(n,k)$ as narrow as possible—we've already qualified our conclusion as possibly wrong!

Let us assume for the moment that $k \in (0 \ldots n)$. For any value of p, the function $P(X(n,p) \leq k) = h(k,p)$ strictly decreases from 1 at $p = 0$ to 0 at $p = 1$. By the intermediate value theorem, there is a unique $p_2 \in (0,1)$ such that $P(X(n,p_2) \leq k) = (1-\alpha)/2$ and

$$q \in (p_2, 1] \quad \Rightarrow \quad P(X(n,q) \leq k) < P(X(n,p_2) \leq k) = (1-\alpha)/2.$$

Thus we reject any $p \in [p_2, 1]$.

Similarly, with $k \in (0 \ldots n)$, the function $P(X(n,p) \geq k) = w(k,p)$ strictly increases from 0 at $p = 0$ to 1 at $p = 1$. Again, by the intermediate value theorem, there is a unique $p_1 \in (0,1)$ such that $P(X(n,p_1) \geq k) = (1-\alpha)/2$ and

$$q \in [0, p_1) \quad \Rightarrow \quad P(X(n,q) \geq k) < P(X(n,p_1) \geq k) = (1-\alpha)/2.$$

Therefore, we reject any $p \in [0, p_1]$. Thus we may select $I(n,k) = (p_1, p_2)$ as a $(100 \times \alpha)\%$ confidence interval with p_1 and p_2 in $(0,1)$. This is, in fact, the traditional choice. In most cases, p_1 and p_2 are numerically estimated from the equations

$$P(X(n,p_1) \geq k) = (1-\alpha)/2 \quad \text{and} \quad P(X(n,p_2) \leq k) = (1-\alpha)/2$$

Web sites that do this are available.

The cases of $k = 0$ and $k = n$ are handled similarly. For $k = 0$, the traditional choice is $I(n,k) = [0, p_2]$ because $P(X(n,p) \geq 0) = 1$ for p and solving for p_1 is impossible. For $k = 1$, the traditional choice is $I(n,k) = [p_1, 1]$ because $P(X(n,p) \leq n) = 1$ and solving for p_2 is impossible.

Let me offer a mild criticism of the traditional choice. It seems to be to be a $(100 \times \dfrac{1+\alpha}{2})\%$ confidence interval: it provides a higher level of confidence than advertised (which, in itself, is not bad). It seems more natural to me to replace $(1-\alpha)/2$ with $1-\alpha$ in the equations that define p_1 and p_2 (thus widening $I(n,k)$

a bit). The wider $I(n, k)$ still enjoys the property that, given that a rejected p is indeed correct, the probability of "something like" k is at most $1 - \alpha$.

Proposition 3. *For $k \in (0 \ldots n)$, we have $p_1 < p_2$.*

Proof. We shall prove this by contradiction. Suppose $p_2 \leq p_1$. Choose some p such that $p_2 \leq p \leq p_1$. Since $p \leq p_1$,

$$P(X(n, p) \geq k) \leq P(X(n, p_1) \geq k) = (1 - \alpha)/2$$

Similarly, since $p_2 \leq p$,

$$P(X(n, p) \leq k) \leq P(X(n, p_2) \leq k) = (1 - \alpha)/2$$

Therefore

$$\begin{aligned} 1 &= P(0 \leq X(n, p) \leq n) \\ &\leq P(X(n, p) \geq k) + P(X(n, p) \leq k) \\ &= 2 * (1 - \alpha)/2 \\ &= 1 - \alpha \end{aligned}$$

and thus $\alpha \leq 0$. Since $\alpha \in (0, 1)$, this contradiction completes the proof. □

References

[Larson82] Larson, Harold J. *Introduction to Probability Theory and Statistical Inference,3rd Edition*, John Wiley & Sons, New York, 1982, pages 390 — 382, especially Theorem 7.3.4.