CONVEXITY OF THE BINOMIAL DISTRIBUTION

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

Abstract. We describe the shape of the Binomial distribution, especially its convexity.

1. Introduction

In this note we describe the shape of the Binomial distribution, especially its convexity. Let $X(n,p)$ denote 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. The Basic Hill Shape

Proposition 1. *For $p \in (0,1)$ let $m = \lceil (n+1)p - 1 \rceil$.*

(1) *For $k \in [0 \ldots n)$,*

$$P(X(n,p) = k+1) > P(X = k) \quad \Leftrightarrow \quad (n+1)p - 1 > k$$

(2) *For $k \in [0 \ldots n)$,*

$$P(X(n,p) = k+1) < P(X = k) \quad \Leftrightarrow \quad (n+1)p - 1 < k$$

(3) *For $k \in [0 \ldots n)$,*

$$P(X(n,p) = k+1) = P(X(n,p) = k) \quad \Leftrightarrow \quad (n+1)p - 1 = k$$

(4) *For r and s in $[0 \ldots n]$,*

$$r < s \le m \quad \textit{implies} \quad P(X = r) < P(X = s)$$

and

$$(n+1)p - 1 < r < s \quad \textit{implies} \quad P(X = r) > P(X = s).$$

(5) *$P(X = k)$ is maximized at $k = m$ and this maximum is unique except when $(n+1)p$ is an integer. When $(n+1)p$ is an integer, the maximum occurs at both $(n+1)p - 1$ and $(n+1)p$ (and nowhere else).*

Date: March 31, 2005.

(6) *For* $k \in [0 \dots n]$, $P(X = k)$ *is minimized at* $k = 0$ *or* $k = n$. *For* $p = 1/2$, *both* $k = 0$ *and* $k = n$ *minimize* $P(X = k)$; *the minimum value is* $(1/2)^n$. *For* $0 < p < 1/2$, $P(X = k)$ *is minimized with* $k = n$ *with minimum value* p^n. *For* $1/2 < p < 1$, $P(X = k)$ *is minimized with* $k = 0$ *with minimum value* $(1 - p)^n$.

Remark. The dual maximum case occurs for n distinct values of p, $p = s/(n+1)$ for integers $1 \leq s \leq n$.

Proof. Let us abbreviate $X(n, p)$ as simply X. Since $p \in (0, 1)$, $P(X = k) > 0$ for all $k \in [0 \dots n]$. For $0 \leq k < n$,

$$\begin{aligned} \frac{P(X = k + 1)}{P(X = k)} &= \frac{n!(n - k)!k!p^{k+1}(1 - p)^{n-k-1}}{n!(k + 1)!(n - k - 1)!p^k(1 - p)^{n-k}} \\ &= \frac{(n - k)p}{(k + 1)(1 - p)}. \end{aligned}$$

The difference $(n - k)p - (k + 1)(1 - p) = (n + 1)p - 1 - k$ and hence

(1) $P(X = k + 1) > P(X = k)$ if and only if $(n + 1)p - 1 > k$.
(2) $P(X = k + 1) < P(X = k)$ if and only if $(n + 1)p - 1 < k$.
(3) $P(X = k + 1) = P(X = k)$ if and only if $(n + 1)p - 1 = k$.

The rest of the proposition now follows quickly. Note that, with $0 < p < 1$, we have $(n+1)p-1 > -1$ and $(n+1)p-1 < (n+1)-1 = n$. Therefore $m = \lceil (n+1)p-1 \rceil$ is in $[0 \dots n]$ and clearly maximizes $P(X = k)$. Since $(n + 1)p$ is in $(0, n + 1)$, when $(n+1)p$ is an integer it must be in $[1, n]$ and thus both $(n+1)p$ and $(n+1)p-1$ are in $[0 \dots n]$. By the previous paragraph, $P(X = k)$ is maximized at both $(n + 1)p$ and $(n + 1)p - 1$.

In the sentence about $r < s \leq m$, because r is an integer one must have $r < (n+1)p-1$ by the definition of $m = \lceil (n+1)p-1 \rceil$; hence $P(X = r) < P(X = r+1)$. This strict inequality continues for each $t \in (r \dots s-1]$: $P(X = t) < P(X = t+1)$.

To prove the last sentence, let $P(X = k)$ be minimized at $k = r \in [0 \dots n]$. Suppose $r > (n + 1)p - 1$. If $r < n$ we would have $P(r) > P(n)$ by Item 4; thus we must have $r = n$. If $r \leq (n + 1)p - 1$, then $r \leq m$. If $r > 0$ we would have $P(r) > P(0)$ by Item 4; thus we must have $r = 0$. By comparing $P(X = 0) = (1 - p)^n$ with $P(X = n) = p^n$, we can determine the minimum value and its location. □

Next we explore the relation between $\lceil (n+1)p-1 \rceil$ and the mean np of $X(n, p)$.

Proposition 2. *Let* $p \in (0, 1)$ *and* n *a positive integer. Abbreviate* $X(n, p)$ *as simply* X. *Note that*

$$-1 < np - 1 < (n + 1)p - 1 < np < (n + 1)p < n + 1$$

If $(n + 1)p$ *is not an integer, then*

(1) $\lceil (n + 1)(1 - p) - 1 \rceil = n - \lceil (n + 1)p - 1 \rceil$.
(2) $np < \lceil (n + 1)p - 1 \rceil$ *if and only if* $n(1 - p) > \lceil (n + 1)(1 - p) - 1 \rceil$.
(3) $np > \lceil (n + 1)p - 1 \rceil$ *if and only if* $n(1 - p) < \lceil (n + 1)(1 - p) - 1 \rceil$.
(4) $np = \lceil (n + 1)p - 1 \rceil$ *if and only if* $n(1 - p) = \lceil (n + 1)(1 - p) - 1 \rceil$.

If $(n + 1)p$ *is an integer, then*

(1) $\lceil (n + 1)(1 - p) - 1 \rceil = (n - 1) - \lceil (n + 1)p - 1 \rceil$.

(2) $\lceil (n+1)p - 1 \rceil < np < \lceil (n+1)p \rceil$. *Thus, by the previous proposition, np is strictly between the two integers k at which $P(X = k)$ is maximized.*
(3) $(n+1)(1-p)$ *is also an integer and the previous two items hold with p and $1-p$ replacing each other (since $1-(1-p) = p$).*

Finally, we mention one more special case: if np is an integer, then $\lceil (n+1)p - 1 \rceil = np$.

Proof. With $p \in (0, 1)$ we have

$$(n+1)p - 1 = np + p - 1 < np + 1 - 1 = np < np + p = (n+1)p < n + 1$$

Also

$$(n+1)p - 1 = np + p - 1 > np - 1 > -1$$

Next, suppose that $(n+1)p$ is not an integer. $m = \lceil (n+1)p - 1 \rceil$. Then $(n+1)p - 1$ is not an integer and therefore $m - 1 < (n+1)p - 1 < m$. Note that

$$\begin{gathered} m - 1 < (n+1)p - 1 < m \\ \Updownarrow \\ -m + 1 > -(n+1)p + 1 > -m \\ \Updownarrow \\ n - m + 1 > n - (n+1)p + 1 > n - m \\ \Updownarrow \\ n - m > (n+1)(1-p) - 1 > n - m - 1 \end{gathered}$$

Therefore $\lceil (n+1)(1-p) - 1 \rceil = n - m$.

Continue with $(n+1)p$ not being an integer. Then

$$\begin{gathered} np < \lceil (n+1)p - 1 \rceil \\ \Updownarrow \\ -np > -\lceil (n+1)p - 1 \rceil \\ \Updownarrow \\ n - np > n - \lceil (n+1)p - 1 \rceil \\ \Updownarrow \\ n(1-p) > \lceil (n+1)(1-p) - 1 \rceil \end{gathered}$$

(where we use what was proved in the previous paragraph). This same argument works equally well with each $<$ and $>$ replaced by its opposite, and with all of them replaced by $=$.

Next, suppose that $(n+1)p$ is an integer. Then $(n+1)p-1$ and $(n+1)(1-p)-1$ are also integers and $m = (n+1)p - 1$. Hence

$$\lceil (n+1)(1-p) - 1 \rceil = (n+1)(1-p) - 1 = n - (n+1)p = n - (m+1) = (n-1) - m$$

By the first item,

$$\lceil (n+1)p - 1 \rceil = (n+1)p - 1 < np < (n+1)p = \lceil (n+1)p \rceil$$

Finally, assume that np is an integer. By the first item $np - 1 < (n+1)p - 1 < np$. Therefore $\lceil (n+1)p - 1 \rceil = np$. □

3. Convexity

We explore the convexity of $X(n,p)$ by examining the change in the forward differences (these are made explicit in the next proposition). Since we are comparing forward differences, we need $n \geq 2$ so that $(0 \ldots n)$ is not empty and we therefore have some changes in forward differences to compare.

Proposition 3. *Let $n \geq 2$ be an integer and $p \in (0,1)$. For k in $[0 \ldots n)$ let $\Delta_k(p)$ be the forward difference of binomial probabilities; specifically,*

$$\Delta_k(p) = P(X(n,p) = k+1) - P(X(n,p) = k)$$

For each k in $(0 \ldots n)$, let

$$p_{k,1} = \frac{(n+1)(k+1) - \sqrt{(k+1)(n+1)(n+1-k)}}{(n+1)(n+2)}$$

and

$$p_{k,2} = \frac{(n+1)(k+1) - \sqrt{(k+1)(n+1)(n+1-k)}}{(n+1)(n+2)}$$

For each k in $(0 \ldots n)$, $p_{k,1}$ and $p_{k,2}$ are in $(0,1)$. Moreover,

1. *The binomial distribution $X(n,p)$ is concave up at k for $p \in (0, p_{k,1})$ and $p \in (p_{k,2}, 1)$ in the sense $\Delta_k(p) - \Delta_{k-1}(p) > 0$ for such p (forward differences are increasing).*
2. *The binomial distribution $X(n,p)$ is concave down at k for $p \in (p_{k,1}, p_{k,2})$ in the sense $\Delta_k(p) - \Delta_{k-1}(p) < 0$ for such p (forward differences are decreasing).*
3. *$\Delta_k(p) - \Delta_{k-1}(p) = 0$ if and only if $p = p_{k,1}$ or $p = p_{k,2}$*

Proof. Let k be in $(0 \ldots n)$. Then

$$\begin{aligned}
&\Delta_k(p) - \Delta_{k-1}(p) \\
&= \tbinom{n}{k+1} p^{k+1}(1-p)^{n-k-1} - \tbinom{n}{k} p^k (1-p)^{n-k} \\
&\quad - \tbinom{n}{k} p^k (1-p)^{n-k} + \tbinom{n}{k-1} p^{k-1}(1-p)^{n-k+1} \\
&= p^{k-1}(1-p)^{n-k-1} \cdot \\
&\quad \left[\tbinom{n}{k+1} p^2 + \tbinom{n}{k-1}(1-p)^2 - 2\tbinom{n}{k} p(1-p)\right] \\
&= \frac{n! p^{k-1}(1-p)^{n-k-1}}{(k+1)!(n-k+1)!} \cdot \\
&\quad \left[(n-k+1)(n-k)p^2 + k(k+1)(1-p)^2 - 2(k+1)(n-k+1)p(1-p)\right] \\
&= \frac{n! p^{k-1}(1-p)^{n-k-1}}{(k+1)!(n-k+1)!} \cdot \\
&\quad \left\{(n+1)(n+2)p^2 - 2(k+1)(n+1)p + k(k+1)\right\}
\end{aligned}$$

Let

$$\Gamma(k,p) = (n+1)(n+2)p^2 - 2(k+1)(n+1)p + k(k+1).$$

Because $0 < p < 1$, $\Delta_k(p) - \Delta_{k-1}(p)$ has the same sign as $\Gamma(k,p)$. We have

$$\frac{\partial \Gamma(k,p)}{\partial p} = 2(n+1)(n+2)p - 2(k+1)(n+1) = 2(n+1)\left[(n+2)p - (k+1)\right].$$

Thus $\Gamma(k,p)$ is a parabola in p opening upward, with vertex at $p = (k+1)/(n+2)$ and

$$\Gamma(k,(k+1)/(n+2)) = -\frac{(k+1)^2(n+1)}{n+2} + k(k+1) = -(k+1)\frac{n+1-k}{n+2} < 0.$$

Note that $\Gamma(k,0) = k(k+1) > 0$ and $\Gamma(k,1) = (n-k)(n+1-k) > 0$. Thus, for each k, there are exactly two values of p, say $p_{k,1} < p_{k,2}$ in $(0,1)$ such that

(1) $\Gamma(k, p_{k,1}) = 0$ and $\Gamma(k, p_{k,2}) = 0$.
(2) $\Gamma(k,p) > 0$ for $p \in [0, p_{k,1})$ and for $p \in (p_{k,2}, 1]$.
(3) $\Gamma(k,p) < 0$ for $p \in (p_{k,1}, p_{k,2})$.

By the quadratic formula, the values of $p_{k,1}$ and $p_{k,2}$ are

$$p_{k,1} = \frac{(n+1)(k+1) - \sqrt{(k+1)(n+1)(n+1-k)}}{(n+1)(n+2)}$$

and

$$p_{k,2} = \frac{(n+1)(k+1) - \sqrt{(k+1)(n+1)(n+1-k)}}{(n+1)(n+2)}$$

□

Proposition 4. *Let $n \geq 2$ be an integer and $p \in (0,1)$. For k in $[0 \ldots n)$ let $\Delta_k(p)$ be the forward difference of binomial probabilities; specifically,*

$$\Delta_k(p) = P(X(n,p) = k+1) - P(X(n,p) = k)$$

For each $p \in (0,1)$, let

$$k_{p,1} = \frac{-1 + 2p(n+1) - \sqrt{1 + 4p(1-p)(n+1)}}{2}$$

and

$$k_{p,2} = \frac{-1 + 2p(n+1) + \sqrt{1 + 4p(1-p)(n+1)}}{2}$$

Then

(1) *We have $\Delta_k(p) - \Delta_{k-1}(p) > 0$ for all k in $(0 \ldots n)$ when*

$$p < \frac{2(n+1) - \sqrt{2n(n+1)}}{(n+1)(n+2)}$$

This last condition is equivalent to $1 > k_{p,2}$.

(2) *We have $\Delta_1(p) - \Delta_0(p) = 0$ and $\Delta_k(p) - \Delta_{k-1}(p) > 0$ for all k in $(1 \ldots n)$ when*

$$p = \frac{2(n+1) - \sqrt{2n(n+1)}}{(n+1)(n+2)}$$

This last condition is equivalent to $1 = k_{p,2}$.

(3) *We have $\Delta_k(p) - \Delta_{k-1}(p) > 0$ for all k in $(0 \ldots n)$ when*

$$p > \frac{n(n+1) + \sqrt{2n(n+1)}}{(n+1)(n+2)}$$

This last condition is equivalent to $k_{p,1} > n-1$.

(4) *We have $\Delta_{n-1}(p) - \Delta_{n-2}(p) = 0$ and $\Delta_k(p) - \Delta_{k-1}(p) > 0$ for all k in $(1 \ldots n-1)$ when*

$$p = \frac{n(n+1) + \sqrt{2n(n+1)}}{(n+1)(n+2)}$$

This last condition is equivalent to $k_{p,1} = n-1$.

(5) *When*

$$\frac{2(n+1)-\sqrt{2n(n+1)}}{(n+1)(n+2)} < p < \frac{n(n+1)+\sqrt{2n(n+1)}}{(n+1)(n+2)}$$

then $\Delta_k(p) - \Delta_{k-1}(p) > 0$ *for* k *in* $(0 \ldots k_{p,1}) \cup (k_{p,2} \ldots n)$; $\Delta_k(p) - \Delta_{k-1}(p) < 0$ *for* k *in* $(k_{p,1} \ldots k_{p,2})$; *and* $\Delta_k(p) - \Delta_{k-1}(p) = 0$ *when* $k = k_{p,1}$ *or* $k = k_{p,2}$.

Proof. We continue with $\Gamma(k,p)$ as in the previous proof. We showed there that, for $p \in (0,1)$ and $k \in (0 \ldots n)$ $\Delta_k(p) - \Delta_{k-1}(p)$ is a positive multiple of

$$\Gamma(k,p) = (n+1)(n+2)p^2 - 2(k+1)(n+1)p + k(k+1)$$

By fixing p and using the quadratic formula to solve for k, we find that $\Gamma(k,p) = 0$ if and only if $k = p_{k,1}$ or $k = k_{p,2}$ where

$$k_{p,1} = \frac{-1+2p(n+1)-\sqrt{1+4p(1-p)(n+1)}}{2}$$

and

$$k_{p,2} = \frac{-1+2p(n+1)+\sqrt{1+4p(1-p)(n+1)}}{2}$$

Note that, with $p \in (0,1)$ and n a positive integer, $1+4p(1-p)(n+1) > 0$ and thus both $k_{p,1}$ and $k_{p,2}$ are real with $k_{p,1} < k_{p,2}$. Since the coefficient of k^2 in $\Gamma(k,p)$ is 1 (and hence positive), $\Gamma(k,p)$ is a parabola in k, opening upward, for a fixed value of p. Therefore $\Gamma(s,p) < 0$ for any real number s in $(k_{p,1}, k_{p,2})$ while $\Gamma(s,p) > 0$ for any real number s in $(-\infty, k_{p,1})$ or in $(k_{p,2}, \infty)$. Thus,

(1) If $1 > k_{p,2}$, we'll have $\Gamma(k,p) > 0$ for all $k \in (0 \ldots n)$.
(2) If $1 = k_{p,2}$, we'll have $\Gamma(1,p) = 0$ while $\Gamma(k,p) > 0$ for $k \in (1 \ldots n)$.
(3) If $n-1 < k_{p,1}$, we'll have $\Gamma(k,p) > 0$ for all $k \in (0 \ldots n)$.
(4) If $n-1 = k_{p,1}$, we'll have $\Gamma(n-1,p) = 0$ while $\Gamma(k,p) > 0$ for $k \in (1 \ldots n-1)$.
(5) The remaining possibility is $n-1 > k_{p,1}$ while $1 < k_{p,2}$. The sign of $\Gamma(k,p)$ may vary with k in this case, as described in the fifth item of this proposition.

To prepare to use the quadratic formula to find p such that $k_{p,2} = 1$, we first simplify the equation:

$$\begin{array}{c} 1 \geq k_{p,2} \\ \Updownarrow \\ 1 \geq \dfrac{-1+2p(n+1)+\sqrt{1+4p(1-p)(n+1)}}{2} \\ \Updownarrow \\ 2 \geq -1+2p(n+1)+\sqrt{1+4p(1-p)(n+1)} \\ \Updownarrow \\ 3-2p(n+1) \geq \sqrt{1+4p(1-p)(n+1)} \\ \Updownarrow \\ (3-2p(n+1))^2 \geq 1+4p(1-p)(n+1) \quad \text{and} \quad 3-2p(n+1) \geq 0 \\ \Updownarrow \\ (n+2)(n+1)p^2-4p(n+1)+2 \geq 0 \quad \text{and} \quad \dfrac{3}{2(n+1)} \geq p \end{array}$$

By the quadratic formula, $(n+2)(n+1)p^2 - 4p(n+1) + 2 = 0$ if and only if

$$p = s = \frac{2(n+1)-\sqrt{2n(n+1)}}{(n+1)(n+2)} \quad \text{or} \quad p = t = \frac{2(n+1)+\sqrt{2n(n+1)}}{(n+1)(n+2)}$$

Because the quadratic $(n+2)(n+1)p^2 - 4p(n+1) + 2$ has a positive coefficient for p^2, we have $(n+2)(n+1)p^2 - 4p(n+1) + 2 < 0$ for p between the two roots above. Therefore,

$$(n+2)(n+1)p^2 - 4p(n+1) + 2 \geq 0 \quad \Leftrightarrow \quad (p \leq s \,\text{or}\, p \geq t)\,.$$

Note that the midpoint between this two solutions s and t is $2/(n+2)$. Note that $2/(n+2) > 3/(2(n+1))$ for $n > 2$ with equality when $n = 2$. Because the two solutions s and t are distinct, the larger solution is bigger than the midpoint and hence, even when $n = 2$, larger than $3/(2(n+1))$. Therefore,

$$p \geq t \quad \Rightarrow \quad p > \frac{3}{2(n+1)}$$

Next, we turn to $p \leq s$. We have

$$
\begin{gathered}
s \leq \frac{3}{2(n+1)} \\
\Updownarrow \\
\frac{2(n+1) - \sqrt{2n(n+1)}}{(n+1)(n+2)} \leq \frac{3}{2(n+1)} \\
\Updownarrow \\
2\left[2(n+1) - \sqrt{2n(n+1)}\right] \leq 3(n+2) \\
\Updownarrow \\
4n + 4 - 3n - 6 \leq 2\sqrt{2n(n+1)} \\
\Updownarrow \\
n - 2 \leq 2\sqrt{2n(n+1)} \\
\Updownarrow \\
(n-2)^2 \leq 4\left[2n(n+1)\right] \quad \text{and} \quad n - 2 \geq 0 \\
\Updownarrow \\
n^2 - 4n + 4 \leq 8n^2 + 8n \quad \text{and} \quad n \geq 2 \\
\Updownarrow \\
0 \leq 7n^2 + 12n - 4 \quad \text{and} \quad n \geq 2 \\
\Updownarrow \\
0 \leq (7n-2)(n+2) \quad \text{and} \quad n \geq 2
\end{gathered}
$$

The quadratic $(7n-2)(n+2)$ has a positive coefficient of n^2 and roots at $2/7$ and -2. Hence $n \geq 2$ implies $(7n-2)(n+2) > 0$ and therefore $s \leq 3/(2(n+1))$. Hence

$$
p \leq s \quad \Rightarrow \quad p \leq \frac{3}{2(n+1)}
$$

It now follows from this paragraph and the previous two paragraphs, that

$$
1 \geq k_{p,2} \quad \Leftrightarrow \quad p \leq s
$$

with equality for $p = s$.

We turn next to finding p such that $n - 1 \leq k_{p,1}$. We begin by noting that

(1) $n - 1 < k_{p,1}$ if and only if $1 > k_{1-p,2}$.
(2) $n - 1 = k_{p,1}$ if and only if $1 = k_{1-p,2}$.
(3) $n - 1 > k_{p,1}$ if and only if $1 < k_{1-p,2}$.
(4) Analogous statements hold hold for $\leq$ and $\geq$.

Here is a proof of the first of these:

$$\begin{aligned}
n-1 &< k_{p,1} \\
&\Updownarrow \\
n-1 &< \frac{-1+2p(n+1)-\sqrt{1+4p(1-p)(n+1)}}{2} \\
&\Updownarrow \\
-1 &< \frac{-2n-1+2p(n+1)-\sqrt{1+4p(1-p)(n+1)}}{2} \\
&\Updownarrow \\
1 &> \frac{-1+2(n+1)-2p(n+1)+\sqrt{1+4p(1-p)(n+1)}}{2} \\
&\Updownarrow \\
1 &> \frac{-1+2(n+1)(1-p)+\sqrt{1+4p(1-p)(n+1)}}{2} \\
&\Updownarrow \\
1 &> k_{1-p,2}
\end{aligned}$$

The analogous inequalities that are listed above can be proved similarly.

To prove the third item of this proposition, by the previous paragraph we know that $n-1 < k_{p,1}$ is equivalent to $1 > k_{1-p,2}$. By the first item, $1 > k_{1-p,2}$ if and only if

$$1-p < \frac{2(n+1)-\sqrt{2n(n+1)}}{(n+1)(n+2)}$$

This is equivalent to

$$1-\frac{2(n+1)-\sqrt{2n(n+1)}}{(n+1)(n+2)} < p$$

Note that

$$\begin{aligned}
1-\frac{2(n+1)-\sqrt{2n(n+1)}}{(n+1)(n+2)} &= \frac{(n+1)(n+2)-2(n+1)+\sqrt{2n(n+1)}}{(n+1)(n+2)} \\
&= \frac{n(n+1)+\sqrt{2n(n+1)}}{(n+1)(n+2)}
\end{aligned}$$

Thus

$$n-1 < k_{p,1} \quad \Leftrightarrow \frac{n(n+1)+\sqrt{2n(n+1)}}{(n+1)(n+2)} < p$$

That proves the equivalence statement that is in Item 3. With $(0\ldots n) \subset (0, k_{p,1})$, we also have $\Delta_k(p) - \Delta_{k-1}(p) > 0$ for all $k \in (0\ldots n)$.

The proof of the fourth item is similar to that of the third item. In the previous paragraph replace every $<$ by $=$ to obtain the equivalence statement that is in Item 4. With $n-1 = k_{p,1}$ we clearly have $\Delta_{n-1}(p) - \Delta_{n-2}(p) = 0$. Also, for $k \in (0\ldots n-1) \subset (0, k_{p,1})$, $\Delta_k(p) - \Delta_{k-1}(p) > 0$. □

For a fixed value of p in $(0,1)$, the random variables $\dfrac{X(n,p)-np}{\sqrt{np(1-p)}}$ converge to the standard normal distribution as $n \to \infty$. For the standard normal distribution, the density function is

$$f(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} \quad \text{with} \quad f''(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}(x^2-1)$$

So f has inflection points at ± 1. After suitable normalization, these are the limits of $k_{p,1}$ and $k_{p,2}$ respectively.

Proposition 5. *Let $p \in (0,1)$ and n a positive integer. Define $k_{p,1}$ and $k_{p,2}$ as in the previous proposition. Then, with p fixed,*

$$\lim_{n\to\infty} \frac{k_{1,p} - np}{\sqrt{np(1-p)}} = -1$$

and

$$\lim_{n\to\infty} \frac{k_{2,p} - np}{\sqrt{np(1-p)}} = -1$$

Proof. We have

$$\frac{-1 + 2p(n+1) \pm \sqrt{1 + 4p(1-p)(n+1)}}{2} - np$$
$$= \frac{-1 + 2p \pm \sqrt{1 + 4p(1-p)(n+1)}}{2}$$

Clearly,

$$\lim_{n\to\infty} \frac{-1 + 2p}{2\sqrt{p(1-p)n}} = 0$$

Also,

$$\frac{\sqrt{1 + 4p(1-p)(n+1)}}{2\sqrt{p(1-p)n}} = \sqrt{\frac{1 + 4p(1-p)(n+1)}{4p(1-p)n}}$$
$$= \sqrt{\frac{1}{4p(1-p)n} + \frac{n+1}{n}}$$
$$\to \sqrt{0+1} = 1 \quad \text{as} \quad n \to \infty$$

It now follows that

$$\lim_{n\to\infty} \frac{k_{1,p} - np}{\sqrt{np(1-p)}} = \lim \frac{-1 + 2p - \sqrt{1 + 4p(1-p)(n+1)}}{2} = -1$$

while

$$\lim_{n\to\infty} \frac{k_{2,p} - np}{\sqrt{np(1-p)}} = \lim \frac{-1 + 2p + \sqrt{1 + 4p(1-p)(n+1)}}{2} = +1$$

□

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.