Statistische Kenngrößen der Binomialverteilung

Hier sollen die Formeln für Erwartungswert und Varianz einer Binomialverteilung bewiesen werden; beide nennt man (neben anderen Größen wie z. B. der Schiefe und Wölbung) „Kenngrößen‘‘ einer Wahrscheinlichkeitsverteilung.

Zunächst verwenden wir der Reihe nach die Definitionen des Erwartungswertes, der Binomialverteilung und des Binomialkoeffizienten; damit können wir schreiben:

$$E(X) = \sum_{k=0}^{n} P_p^n(X = k) \cdot k = \sum_{k=0}^{n} B(n; p; k) \cdot k = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \cdot k$$
$$= \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!}\, p^k (1-p)^{n-k} \cdot k$$

In der Summe trägt aber der Summand mit $k = 0$ gar nichts bei, denn dieser ist (wegen des $\cdot\, k$ hinten) ja einfach gleich Null. Also können wir die Summe auch gleich bei $k = 1$ anfangen lassen,

$$E(X) = \sum_{k=1}^{n} \frac{n!}{k!\,(n-k)!}\, p^k (1-p)^{n-k} \cdot k$$

Wir haben jetzt Summanden, bei denen k bei 1 anfängt und dann durchläuft bis n. Genau dieselben Summanden erhalten wir, wenn wir k von 0 bis $n-1$ schreiben, in der Summe dann aber überall $k+1$ schreiben statt nur k,

$$E(X) = \sum_{k=0}^{n-1} \frac{n!}{(k+1)!\,(n-(k+1))!}\, p^{k+1} (1-p)^{n-(k+1)} \cdot (k+1).$$

Nun nutzen wir einige algebraische Umformungen: Statt $n!$ können wir auch $n \cdot (n-1)!$ schreiben, statt $n-(k+1)$ auch $(n-1)-k$ und statt p^{k+1} auch $p \cdot p^k$. Außerdem können wir auch das $(k+1)!$ im Nenner mit dem Faktor $(k+1)$ ganz hinten kürzen:

$$\frac{(k+1)}{(k+1)!} = \frac{(k+1)}{(k+1)\cdot k!} = \frac{1}{k!}.$$

Damit bleibt schließlich:

$$E(X) = \sum_{k=0}^{n-1} \frac{n \cdot (n-1)!}{k!\,((n-1)-k)!}\, p \cdot p^k (1-p)^{(n-1)-k}.$$

Die Faktoren n und p hängen nicht vom Summationsindex k ab, die können wir also auch vor das Summenzeichen schreiben (d. h. aus der Summe ausklammern):

$$E(X) = n \cdot p \cdot \sum_{k=0}^{n-1} \frac{(n-1)!}{k!\,\big((n-1)-k\big)!}\, p^k (1-p)^{(n-1)-k}.$$

Jetzt noch alles rückwärts: Definitionen des Binomialkoeffizienten und der Binomialverteilung einsetzen,

$$E(X) = n \cdot p \cdot \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{(n-1)-k} = n \cdot p \cdot \sum_{k=0}^{n-1} B(n-1; p; k) = n \cdot p \cdot \sum_{k=0}^{n-1} P_p^{n-1}(X = k).$$

In der übrigen Summe wird nun aber einfach über *alle* möglichen Zufallswerte summiert, also muss diese Summe Eins ergeben. Und damit bleibt tatsächlich

$$\boxed{E(X) = n \cdot p}$$

Die Rechnung für die Varianz geht ganz ähnlich. Dafür nutzen wir den Verschiebungssatz, berechnen also zunächst den Erwartungswert von X^2:

$$E(X^2) = \sum_{k=0}^{n} P_p^n(X = k) \cdot k^2 = \sum_{k=0}^{n} B(n; p; k) \cdot k^2 = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \cdot k^2$$
$$= \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!}\, p^k (1-p)^{n-k} \cdot k^2.$$

Um diese Summe zu vereinfachen, nutzen wir noch einen Trick: Wir schreiben das k^2 um zu

$$k^2 - k + k = k(k-1) + k$$

und teilen die Summe dann auf in zwei Summen,

$$E(X^2) = \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!}\, p^k (1-p)^{n-k} \cdot k \cdot (k-1) + \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!}\, p^k (1-p)^{n-k} \cdot k.$$

Die zweite Summe kennen wir aber schon, das ist ja genau $E(X) = n \cdot p$. Die erste Summe wird nun prinzipiell genauso umgeschrieben, wie wir es vorhin schon gemacht hatten, nur fallen diesmal die ersten beiden Summenglieder weg statt nur des ersten, und dementsprechend wird k dann durch $k+2$ ersetzt statt durch $k+1$:

$$\sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!}\, p^k (1-p)^{n-k} \cdot k \cdot (k-1) = \sum_{k=2}^{n} \frac{n!}{k!\,(n-k)!}\, p^k (1-p)^{n-k} \cdot k \cdot (k-1)$$

$$= \sum_{k=0}^{n-2} \frac{n!}{(k+2)!\,(n-(k+2))!}\, p^{k+2} (1-p)^{n-(k+2)} \cdot (k+1) \cdot k$$

$$= \sum_{k=0}^{n-2} \frac{n \cdot (n-1) \cdot (n-2)!}{k!\,((n-2)-k)!}\, p^2 p^k (1-p)^{(n-2)-k}$$

$$= n \cdot (n-1) \cdot p^2 \cdot \sum_{k=0}^{n-2} \frac{(n-2)!}{k!\,((n-2)-k)!}\, p^k (1-p)^{(n-2)-k}$$

$$= n \cdot (n-1) \cdot p^2 \cdot \sum_{k=0}^{n-2} \binom{n-2}{k} p^k (1-p)^{(n-2)-k} = n \cdot (n-1) \cdot p^2 \cdot \sum_{k=0}^{n-2} B(n-2;p;k)$$

$$= n \cdot (n-1) \cdot p^2 \cdot \sum_{k=0}^{n-2} P_p^{n-2}(X=k) = n \cdot (n-1) \cdot p^2 \cdot 1$$

Daraus folgt

$$E(X^2) = n \cdot (n-1) \cdot p^2 + n \cdot p,$$

und daraus wiederum

$$Var(X) = E(X^2) - \big(E(X)\big)^2 = n \cdot (n-1) \cdot p^2 + n \cdot p - (n \cdot p)^2$$
$$= n^2 \cdot p^2 - n \cdot p^2 + n \cdot p - n^2 \cdot p^2 = n \cdot p - n \cdot p^2,$$

also ist tatsächlich

$$\boxed{Var(X) = n \cdot p \cdot (1-p).}$$