1 Kommentar på uppgift 2.51

I uppgift 2.51 skall man härleda både väntevärde och varians för en binomialfördelad, *Bin(n,p)* variabel Y.

$$\begin{aligned} E[Y] &= np \\ V[Y] &= np(1-p) \end{aligned} \tag{1}$$

Ett enkelt sätt att bevisa påståendet är att utgå från Bernoullifördelade stokastiska variabler x_i där

$$\begin{cases} x_i = 1 \text{ med sannolikhet } p \\ x_i = 0 \text{ med sannolikhet } 1-p \end{cases} \tag{2}$$

Väntevärde och varians beräknas enkelt från sina definitioner:

$$\begin{cases} E[X_i] = 0 \cdot (1-p) + 1 \cdot p = p \\ V[X_i] = E[(X_i - E[X_i])^2] = (0-p)^2 \cdot (1-p) + (1-p)^2 \cdot p = p(1-p) \end{cases} \tag{3}$$

Låt oss nu bilda variabeln Y

$$Y = X_1 + X_2 + \ldots + X_n \tag{4}$$

De vanliga räknereglerna för väntevärde och varians ger då

$$\begin{aligned} E[Y] = E[X_1 + \ldots + X_n] = \sum_{i=1}^n E[X_i] = nE[X_i] = np \\ V[Y] = V[X_1 + \ldots + X_n] = [oberoende] = \sum_{i=1}^n V[X_i] = np(1-p) \end{aligned} \tag{5}$$

Ett annat sätt att härleda dessa formler är att göra som i Montgomerys facit. Det är ganska krångliga formler. Jag härleder uttrycket för väntevärdet. Uttrycket för variansen är krångligare.

$$
\begin{aligned}
E[Y] &= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=1}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x} \\
&= [\text{låt } y = x-1] \\
&= \sum_{y=0}^{n-1} \frac{n!}{y!(n-y-1)!} p^{y+1} (1-p)^{n-y-1} \\
&= np \sum_{y=0}^{n-1} \frac{(n-1)!}{y!(n-1-y)!} p^{y} (1-p)^{n-1-y} \\
&= np \sum_{y=0}^{n-1} \binom{n-1}{y} p^{y} (1-p)^{n-1-y} \\
&= \left[\text{Utnyttja likheten} (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \text{ med } a = p \text{ och } b = 1-p\right] \\
&= np(p + (1-p))^{n-1} = np \cdot 1 = np
\end{aligned}
$$

Den som vill se härledningen för variansen med hjälp av summorna kan t.ex. studera George Casella och Roger Bergers bok ”Statistical Inference”.