

Vi ger här ett bevis för att om  $X$  är  $Hyper(N, n, p)$  så är  $E(X) = np$  och  $V(X) = np(1-p)\frac{N-n}{N-1}$ .  
 Betrakta en urnmodell där vi drar  $n$  kulor på måfå ur en urna utan återläggning där urnan innehåller  $N$  kulor, varav andelen  $p$  är röda. Vi vet då att  $X$ =antalet röda bland de  $n$  är  $Hyper(N, n, p)$ . Vi vill nu ge en representation i form av 0-1-variabler på följande sätt: Låt

$$U_i = \begin{cases} 1 & \text{om kulan i den } i\text{:te omgången är röd} \\ 0 & \text{annars.} \end{cases}$$

Vi inser att  $P(U_i = 1) = p$  och  $P(U_i = 0) = 1 - p$  och att  $X = \sum_{i=1}^n U_i$ .

Notera att  $U_1, U_2, \dots, U_n$  blir svagt beroende på grund av att urnans sammansättning förändras efter de successiva dragningarna. Dock är  $U_i$ :na likafördelade av symmetriskäl – man kan uppfatta det som ett resultat av att man skulle kunna numrera om dragningsomgångarna.

Vi har alltså  $X = \sum_{i=1}^n U_i$  och får lätt  $E(X) = E(\sum_{i=1}^n U_i) = \sum_{i=1}^n E(U_i) = nE(U_1) = np$ . För att få variansen kan man utnyttja att  $V(X) = E(X^2) - (E(X))^2$  och vad som återstår är alltså att beräkna  $E(X^2)$ .

Vi har

$$E(X^2) = E\left(\left(\sum_{i=1}^n U_i\right)^2\right) = E\left(\left(\sum_{i=1}^n U_i\right)\left(\sum_{j=1}^n U_j\right)\right) = E\left(\sum_{i=1}^n \sum_{j=1}^n U_i U_j\right) = \sum_{i=1}^n \sum_{j=1}^n E(U_i U_j)$$

Vi delar nu upp denna summa i de  $n$  termer där  $i = j$  och de  $n(n-1)$  termer där  $i \neq j$ . Vi får då på grund av symmetrin

$$E(X^2) = nE(U_1^2) + n(n-1)E(U_1 U_2) = np + n(n-1)P(U_1 = 1; U_2 = 1).$$

Vi har

$$P(U_1 = 1; U_2 = 1) = P(U_1 = 1)P(U_2 = 1|U_1 = 1) = p\frac{Np-1}{N-1}$$

eftersom om  $U_1 = 1$  (som har sannolikhet  $p$ ) så finns  $Np-1$  st röda att välja på bland de  $N-1$  tillgängliga vid den andra dragningen.

Vi får alltså

$$V(X) = E(X^2) - (E(X))^2 = np + n(n-1)p\frac{Np-1}{N-1} - (np)^2 =$$

som efter förenkling ger

$$V(X) = np(1-p)\frac{N-n}{N-1}.$$