

Binomialfördelningen

Förutsättningar:

1. I ett experiment utför man n identiska försök. Varje försök har två möjliga utfall (Bernoulli-försök), t. ex. lyckas/misslyckas.
2. Sannolikheten att lyckas, p , är konstant från försök till försök (sannolikheten att misslyckas är $q = 1 - p$).
3. De n försöken sker oberoende av varandra.

Då är X = Antalet lyckade försök, en binomialfördelad variabel med parametrarna n och p .

Kortfattat betecknas det

$$X \sim \text{Bin}(n; p) \quad X \text{ är } \text{Bin}(n; p) \quad X \in \text{Bin}(n; p).$$

Sannolikhetsfördelningen för X ges av

$$p(x) = \binom{n}{x} p^x q^{n-x},$$

där x antar alla värden upp till n , dvs $x = 0, 1, 2, \dots, n$.

Dessutom

$$E(X) = np$$

$$V(X) = npq$$

$$SD(X) = \sqrt{npq}$$

Exempel: Två identiska oberoende kast med ett symmetriskt mynt. Låt X = Antal krona. Ge sannolikhetsfördelning, väntevärde och varians för X .

Lösningalternativ 1:

Möjliga fall: $kl, kl \quad kl, kr \quad kr, kl \quad kr, kr$

ger följande sannolikhetsfördelning

x	0	1	2
$\Pr(X = x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

väntevärde

$$E(X) = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1$$

och varians

$$\begin{aligned} V(X) &= (0-1)^2 \cdot \frac{1}{4} + (1-1)^2 \cdot \frac{1}{2} + (2-1)^2 \cdot \frac{1}{4} \\ &= \frac{1}{2}. \end{aligned}$$

Lösningalternativ 2: Identifiera fördelningen för X .

$$X \sim \text{Bin}(2; 0.5)$$

$$\Pr(X=0) = \binom{2}{0} 0.5^0 0.5^2 = 0.25$$

$$\Pr(X=1) = \binom{2}{1} 0.5^1 0.5^1 = 0.5$$

$$\Pr(X=2) = \binom{2}{2} 0.5^2 0.5^0 = 0.25$$

$$E(X) = np = 2 \cdot 0.5 = 1$$

$$V(X) = npq = 2 \cdot 0.5 \cdot 0.5 = 0.5$$

Exempel: $X \sim \text{Bin}(10; 0.2)$

Beräkna

a)

$$\begin{aligned}\Pr(1 \leq X \leq 3) &= \Pr(X \leq 3) - \Pr(X < 1) \\ &= \Pr(X \leq 3) - \Pr(X \leq 0) \\ &= 0.8791 - 0.1074 = 0.7717\end{aligned}$$

b)

$$\Pr(X \geq 7) = 1 - \Pr(X \leq 6) = 1 - 0.9991 = 0.0009$$

c)

$$\begin{aligned}\Pr(X = 5) &= \Pr(X \leq 5) - \Pr(X \leq 4) \\ &= 0.9936 - 0.9672 = 0.0264\end{aligned}$$

Poissonfördelningen

Ibland har man experiment som resulterar i att man räknar antal gånger en speciell händelse inträffar inom en given tidsperiod eller på ett givet fysiskt objekt. Händelserna inträffar oberoende av varandra med viss intensitet μ .

Exempel:

- Antal telefonsamtal till en växel mellan 9.00 och 10.00
- Antal tryckfel på en slumpmässigt vald sida i en bok
- Antal kunder som kommer till en affär mellan 12.00 och 14.00

Varje frekvens kan betraktas som en stokastisk variabel som är poissonfördelad med parameter μ , dvs

$$X \sim \text{Po}(\mu)$$

Sannolikhetsfördelningen ges av

$$\Pr(X = x) = \frac{\mu^x e^{-\mu}}{x!}, \text{ för } x = 0, 1, 2, \dots$$

och väntevärde och varians av

$$E(X) = \mu$$

$$V(X) = \mu$$

Exempel:

Antalet flygkrascher i civilflyget i Sverige är i genomsnitt 2 per år. Den s.v. X = Antalet flygkrascher i civilflyget under ett slumpmässigt valt år är då $\text{Po}(2)$. Beräkna $\Pr(X \leq 1)$.

Alternativ 1:

$$\begin{aligned}\Pr(X \leq 1) &= \Pr(X = 0) + \Pr(X = 1) \\ &= \frac{2^0 e^{-2}}{0!} + \frac{2^1 e^{-2}}{1!} \\ &= 0.135 + 0.270 = 0.406\end{aligned}$$

Alternativ 2:

$$\Pr(X \leq 1) = 0.406 \text{ enligt tabell 2.}$$

Exempel 10 (Körner sid 87):

0.1 % av befolkningen dör varje år i en viss sjukdom. Slumpmässigt väljs 1000 personer ur befolkningen.

Vad är sannolikheten att minst två personer i stickprovet dör i sjukdomen inom ett år?

X = Antalet som dör pga sjukdomen inom ett år.

$$X \sim \text{Bin}(1000; 0.001)$$

$$\begin{aligned}\Pr(X \geq 2) &= 1 - \Pr(X \leq 1) \\ &= 1 - \Pr(X = 0) - \Pr(X = 1) \\ &= 1 - 0.999^{1000} \\ &\quad - \binom{1000}{1} 0.001^1 0.999^{999} \\ &= 1 - 0.3677 - 0.3681 = 0.2642\end{aligned}$$

Om n är stort och p är litet (en tumregel är $n \geq 20$ och $p \leq 0.05$) kan binomialfördelningen approximeras med en poissonfördelning. Vi behöver väntevärdet

$$E(X) = \mu = np = 1000 \cdot 0.001 = 1$$

$$\begin{aligned}\Pr(X \geq 2) &= 1 - \Pr(X \leq 1) = 1 - 0.7358 \\ &= 0.2642 \text{ enligt tabell 2.}\end{aligned}$$

Hypergeometrisk fördelning (jfr med binomialfördelningen)

I en ändlig (liten) population av storlek N dras ett slumpmässigt stickprov utan återläggning av storlek n . Om p är andelen individer i populationen som har en viss egenskap är Np antalet individer som har egenskapen, medan resten $N - Np$ saknar egenskapen. Om X = "Antalet i stickprovet som har egenskapen" gäller att sannolikhetsfördelningen ges av

$$\Pr(X = x) = \frac{\binom{Np}{x} \binom{N-Np}{n-x}}{\binom{N}{n}},$$

för $x < n, x < Np, n - x < N - Np$

Vi säger att X är hypergeometriskt fördelad med parametrar N, n och p . Väntevärde och varians ges av

$$\begin{aligned}E(X) &= np \\ V(X) &= npq \frac{N-n}{N-1}\end{aligned}$$

Exempel:

I en låda ligger tio lampor. 30 % av lamporna är defekta. Man drar fem stycken utan återläggning.

Vad är sannolikheten att högst en av de dragna lamporna är defekta?

Låt X =Antal defekta lampor i urvalet. Då gäller att

$$\begin{aligned}\Pr(X \leq 1) &= \Pr(X = 0) + \Pr(X = 1) \\ &= \frac{\binom{3}{0} \binom{10-3}{5-0}}{\binom{10}{5}} + \frac{\binom{3}{1} \binom{10-3}{5-1}}{\binom{10}{5}} \\ &= \frac{56}{252} = 0.222.\end{aligned}$$

Om N är stort i förhållande till n (Tumregel: $\frac{n}{N} < 0.1$) gäller approximativt att

X är $Bin(n; p)$

$$E(X) = np$$

$$V(X) \approx npq$$

eftersom

$$\frac{N-n}{N-1} = \frac{N}{N-1} - \frac{n}{N-1} \approx 1 - 0 = 1.$$