Stichproben mit seltener Merkmalsausprägung

Untersucht werde ein Merkmal mit der Trefferwahrscheinlichkeit $p = 0{,}004$. Man nimmt eine Stichprobe vom Umfang $n = 2000$. Wie groß ist die Wahrscheinlichkeit, daß das Merkmal in der Stichprobe 50-mal auftaucht?

$$\binom{2000}{50} = 199616957342792362110329035972627812510043302235611927600885824928160$$

$$(0{,}004)^{50} = 1{,}26765 \cdot 10^{-120}$$

$$(0{,}996)^{1950} = 0{,}000403376$$

$$\mathbf{P(\ X = 50) = B(\ 2000\ ;\ 0{,}004;\ 50\ )} = \binom{2000}{50} \cdot 0{,}004^{50} \cdot 0{,}996^{1950} = 1{,}02072 \cdot 10^{-23}$$

Zentrales rechnerisches Problem: **Produkt <u>sehr großer</u> und <u>sehr kleiner</u> Zahlen** !

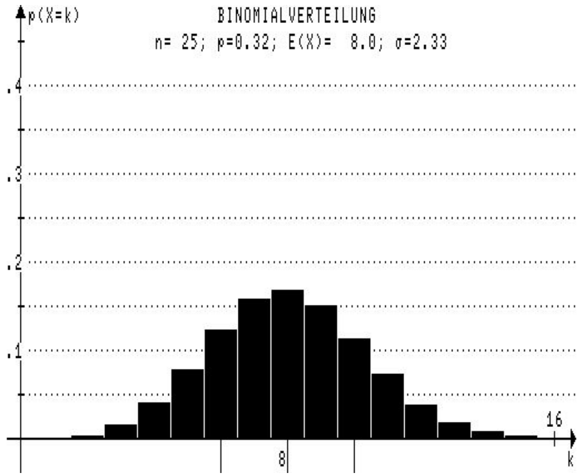
Es gilt: $\boldsymbol{\mu = E(X) = 8}$!

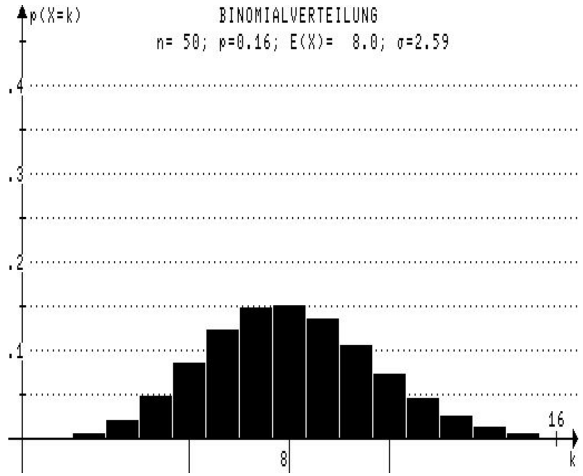
n	25	50	75	100	125	150
p	0,320	0,160	$0{,}10\overline{6}$	0,080	0,064	$0{,}05\overline{3}$

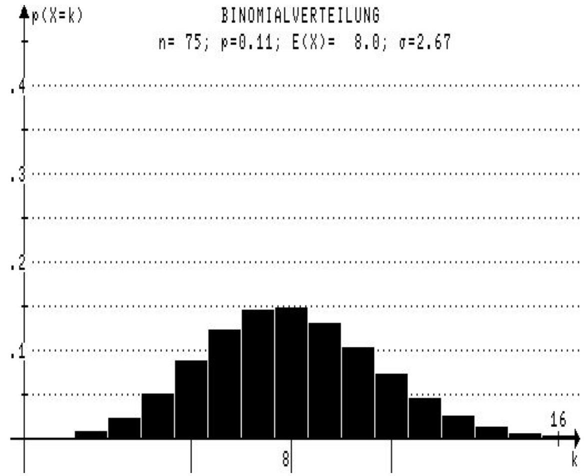
Wie verändern sich Binomialverteilungen für größeres **n** und kleineres **p** (bei festem Produkt $n \cdot p = \mu$) ?

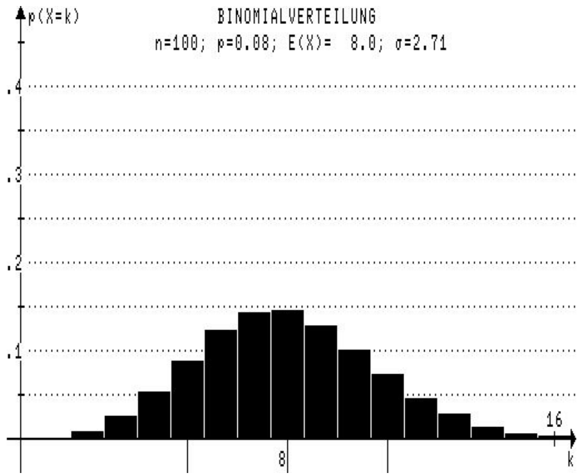
Stichproben mit seltener Merkmalsausprägung

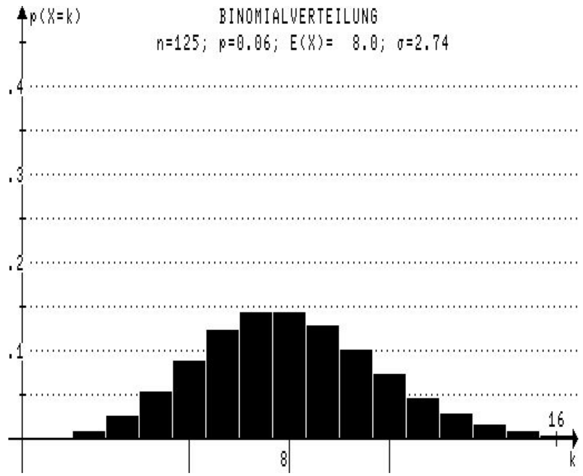
Es gilt stets: **μ = E(X) = 8** ! - Wie verändern sich die Binomialverteilungen für größeres **n** und kleineres **p** ?[1]

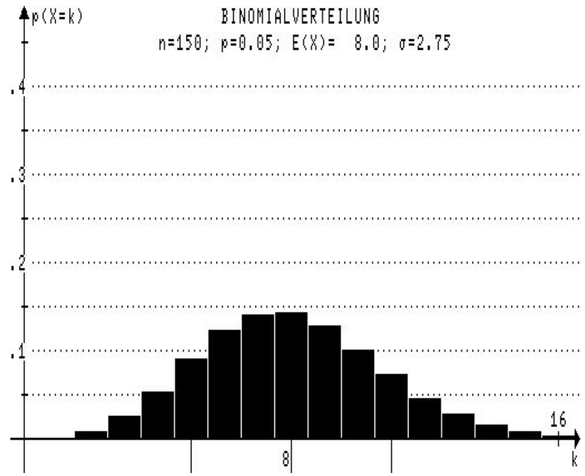












[1] <u>Beachte:</u> In den Graphiken ist die angezeigte Trefferwahrscheinlichkeit gerundet. - Was wären die exakten Trefferwahrscheinlichkeiten unter der Voraussetzung, dass n und μ korrekt angezeigt werden?

Stichproben mit seltener Merkmalsausprägung

P(X = 0) =	**B(2000 ; 0,004; 0)** ≈	0,000330124
P(X = 1) =	**B(2000 ; 0,004; 1)** ≈	0,0026516
P(X = 2) =	**B(2000 ; 0,004; 2)** ≈	0,0106437
P(X = 3) =	**B(2000 ; 0,004; 3)** ≈	0,0284686
P(X = 4) =	**B(2000 ; 0,004; 4)** ≈	0,0570801
P(X = 5) =	**B(2000 ; 0,004; 5)** ≈	0,0915115
P(X = 6) =	**B(2000 ; 0,004; 6)** ≈	0,122199
P(X = 7) =	**B(2000 ; 0,004; 7)** ≈	0,139796
P(X = 8) =	**B(2000 ; 0,004; 8)** ≈	0,139867
P(X = 9) =	**B(2000 ; 0,004; 9)** ≈	0,124326
P(X = 10) =	**B(2000 ; 0,004; 10)** ≈	0,0994107
P(X = 11) =	**B(2000 ; 0,004; 11)** ≈	0,0722261
P(X = 12) =	**B(2000 ; 0,004; 12)** ≈	0,0480782
P(X = 13) =	**B(2000 ; 0,004; 13)** ≈	0,0295272
P(X = 14) =	**B(2000 ; 0,004; 14)** ≈	0,0168303
P(X = 15) =	**B(2000 ; 0,004; 15)** ≈	0,00894914
P(X = 16) =	**B(2000 ; 0,004; 16)** ≈	0,00445885

Stichproben mit seltener Merkmalsausprägung

Bei festem Produkt $n \cdot p = E(X) = \mu$ gilt: $p = \dfrac{\mu}{n}$

$$B(n;p;k) = \binom{n}{k} \cdot \left(\frac{\mu}{n}\right)^k \cdot \left(1 - \frac{\mu}{n}\right)^{n-k}$$

$$= \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} \cdot \frac{\mu^k}{n^k} \cdot \left(1 - \frac{\mu}{n}\right)^n \cdot \left(1 - \frac{\mu}{n}\right)^{-k}$$

$$= \frac{n \cdot (n-1) \cdots (n-k+1)}{n^k} \cdot \frac{\mu^k}{k!} \cdot \left(\left(1 + \frac{(-\mu)}{n}\right)^{\frac{n}{(-\mu)}}\right)^{-\mu} \cdot \left(1 - \frac{\mu}{n}\right)^{-k}$$

$$= \frac{n}{n} \cdot \frac{n-1}{n} \cdot \ldots \cdot \frac{n-k+1}{n} \cdot \frac{\mu^k}{k!} \cdot \left(\left(1 + \frac{1}{\frac{n}{(-\mu)}}\right)^{\frac{n}{(-\mu)}}\right)^{-\mu} \cdot \left(1 - \frac{\mu}{n}\right)^{-k}$$

$$\xrightarrow[n \to \infty]{} \quad 1 \cdot \frac{\mu^k}{k!} \cdot e^{-\mu} \cdot 1$$

$$\approx \frac{\mu^k}{k!} \cdot e^{-\mu}$$

Poisson - Näherung der Binomialverteilung