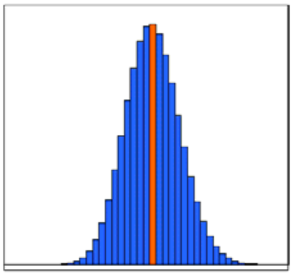
Binomialfördelningen



Med binomialfördelningen beräknar man:

Slh. = Sannolikhet

- Slh. att få 5 ggr. krona i 10 myntkast
- Slh. att få 10 femmor i 100 tärningskast
- Slh. att få 5 vita utav 20 dragna kulor *(ur en urna med många svarta och vita kulor)*
- ... och mycket mer

Bernoulli-Experiment

Johann Bernoulli

Slh. = Sannolikhet

- Ett slumpförsök har 2 alternativa utfall: A, A*
- Vi känner slh:a för båda utfallen: P(A)=*p och P(A*)=1-p*
  - mynt: P(*krona*) = p = 0.5 och P(*vapen*) = 1 - p = 0.5
  - tärning: P(femma) = 1/6 och P(ingen femma)=1 – p = 5/6
- Vi upprepar detta försök flera (n) gånger.
- *Att beräkna:* slh. att det blir k gånger utfall A
  - slh. att få 5 ggr. krona i 10 myntkast (n=10 k=5 p=0.5)
  - slh. att få 10 femmor i 100 tärningskast (n=100 k=10 p=1/6)

Bernoulli-Experiment

utfall	A	A*
slh.	p	1 - p
antal	k	n - k

Slumpvariabel X: antalet utfall A

$\Omega_X=\{ 0, 1, 2, ... n \}$

$p_X(k)$ = ???

Det här vill man helst veta för alla k som kan förekomma, dvs. för hela utfallsrummet Ω

OBS !! Blanda inte ihop:

- p är slh. att lyckas i ett enda försök
- $p_X(k)$ är slh. att lyckas k gånger i n försök

Exempel: 5 myntkast (n=5)

$$P_X(k=5)=\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{32}=0.03125$$

s.v. X = antal framsidor

$$P_X(k=4)=5\cdot\frac{1}{32}=0.15625$$

osv. ...

Exempel: 5 myntkast (n=5)

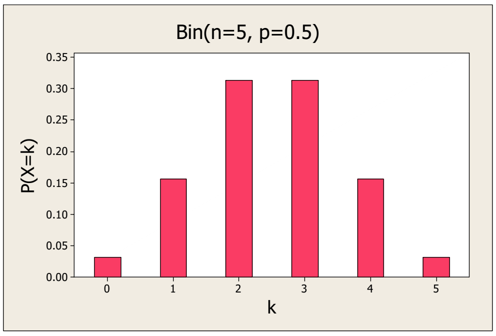
$$P_X(k=3)=\binom{5}{3}\cdot p^3\cdot(1-p)^2=10\cdot\frac{1}{2^3}\cdot\frac{1}{2^2}=\frac{10}{32}=0.3125$$

$$P_X(k=2)=\binom{5}{2}\cdot p^2\cdot(1-p)^3=10\cdot\frac{1}{2^2}\cdot\frac{1}{2^3}=\frac{10}{32}=0.3125$$

$$P_X(k=1)=\binom{5}{1}\cdot p^1\cdot(1-p)^4=5\cdot\frac{1}{2^1}\cdot\frac{1}{2^4}=\frac{5}{32}=0.15625$$

$$P_X(k=0)=\binom{5}{0}\cdot p^0\cdot(1-p)^5=1\cdot 1\cdot\frac{1}{2^2}=\frac{1}{32}=0.03125$$

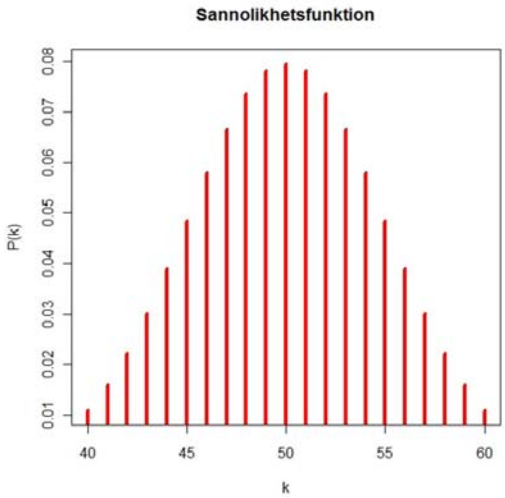
Sannolikhetsfunktion (5 myntkast)



Binomial.MPJ

Calc / Prob. Distrib. / Binomial Probability
Number of trials: 5
Event probability: 0.5
Input column: C1
Optional storage: C2

x	P(X = k)
0	0.03125
1	0.15625
2	0.31250
3	0.31250
4	0.15625
5	0.03125



Allmän beräkning av Sannolikhetsfunktionen

utfall	A	A*
slh.	p	1 - p
antal	k	n - k

k = 0,1, 2, ... n

$$X \in Bin(n,p)$$

$$P_X(k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

En slumpvariabel X med ovanstående sannolikhetsfunktion kallas binomialfördelad.
Man skriver X ~ Bin(n,p) *n = antalet försök* *p = slh. ”att lycka” (A kommer upp)*

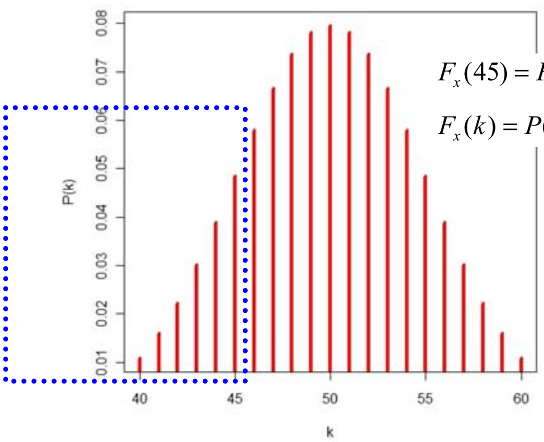
Fördelningsfunktionen för Bin(n,p)

- Sannoliheten att k blir lika med eller mindre än något tal:
  - slh. att få ≤ 5 ggr. krona i 10 myntkast (dvs. max. 5 ggr. krona)
  - slh. att få max. 10 femmor i 100 tärningskast

$$P(max\ 3\ femmor) = P_X(0) + P_X(1) + P_X(2) + P_X(3)$$

”eller blir +”

Fördelningsfunktion för k = 45



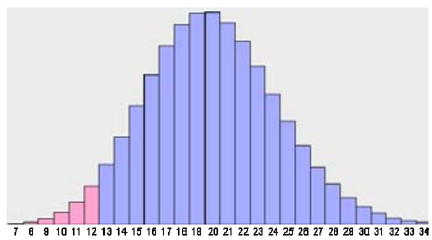
$$F_x(45) = P(X \le 45)$$

$$F_x(k) = P(X \le k)$$

Fördelningsfunktionen

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad k = 0, 1, \dots n$$ pdf

$$F_X(k) = \sum_{j=0}^{k} \binom{n}{j} p^j (1-p)^{n-j}$$ cdf



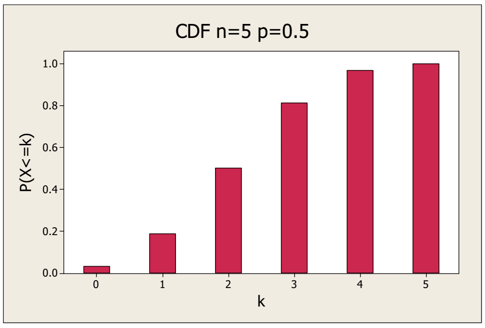
Tabell över fördelningsfunktionen för binomialfördelningen

Tabell 6. Binomialfördelningen k

$P(X \le x)$ där $X \in \text{Bin}(n, p)$.
För $p > .5$ utnyttja att $P(X \le x) = P(Y \ge n - x)$ där $Y \in \text{Bin}(n, 1 - p)$

n	x \ p	0.05	0.10	0.15	0.20	0.25	0.30	0.40	0.50
2	0	.90250	.81000	.72250	.64000	.56250	.49000	.36000	.25000
	1	.99750	.99000	.97750	.96000	.93750	.91000	.84000	.75000
3	0	.85737	.72900	.61412	.51200	.42188	.34300	.21600	.12500
	1	.99275	.97200	.93925	.89600	.84375	.78400	.64800	.50000
	2	.99987	.99900	.99662	.99200	.98438	.97300	.93600	.87500
4	0	.81451	.65610	.52201	.40960	.31641	.24010	.12960	.06250
	1	.98598	.94770	.89048	.81920	.73828	.65170	.47520	.31250
	2	.99952	.99630	.98802	.97280	.94922	.91630	.82080	.68750
	3	.99999	.99990	.99949	.99840	.99609	.99190	.97440	.93750
5	0	.77378	.59049	.44371	.32768	.23730	.16807	.07776	.03125
	1	.97741	.91854	.83521	.73728	.63281	.52822	.33696	.18750

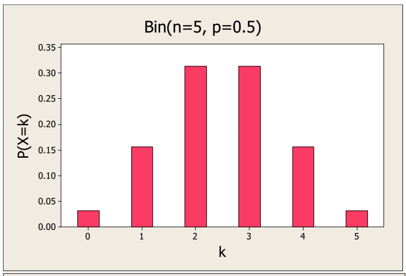
Fördelningsfunktion (5 myntkast)



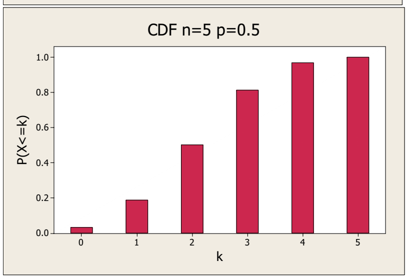
Calc / Prob. Distrib. / Binomial
Cumulative Probability
Number of trials: 5
Event probability: 0.5
Input column: C1
Optional storage: C2

x	P(X <= x)
0	0.03125
1	0.18750
2	0.50000
3	0.81250
4	0.96875
5	1.00000

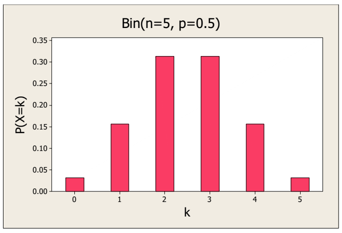
Sannolikhets-funktion

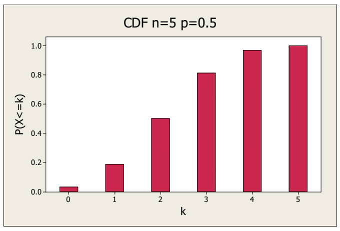


Fördelnings-funktion



Beräkning av en sannolikhet med hjälp av fördelningsfunktionen

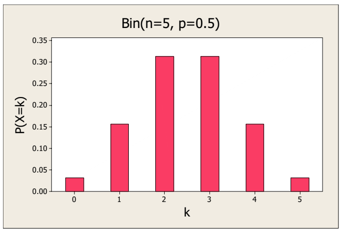


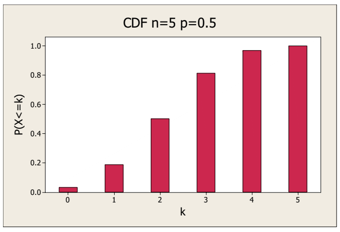


$$p_X(4) = F_X(4) - F_X(3)$$

$$p_X(k) = F_X(k) - F_X(k-1)$$

Beräkning av en sannolikhet med hjälp av fördelningsfunktionen





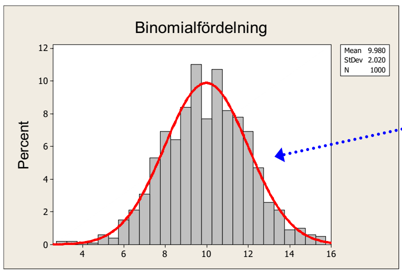
$$P(2 < X \le 4) = F_X(4) - F_X(2)$$

$$P(a < X \le b) = F_X(b) - F_X(a)$$

Slumpvariabeln X är diskret → det är mycket viktigt att använda likhetstecknen korrekt.

Normalapproximation för Bin

Låt X vara en binomialfördelad slumpvariabel.
Om många försök görs (dvs. om n är stor) → X är ungefär normalfördelad:



Den röda kurvan är täthetsfunktionen för en normalfördelning

Sats från Moivre-Laplace

Normalapproximation för Bin

Låt X vara en binomialfördelad slumpvariabel. Många försök görs (dvs. n är stor) → X är ungefär normalfördelad ... med binomialfördelningens väntevärde och varians.

$$X \in Bin(n,p) \quad \text{med } n \cdot p \cdot (1-p) \geq 10$$
$$\Longrightarrow X \in N\left(np, \sqrt{np(1-p)}\right)$$

Sats från Moivre-Laplace


Tilläggsinformation

För stora n är X ungefär normalfördelad med binomialfördelningens väntevärde och standardavvikelse:

$$X \in N\left(np, \sqrt{np(1-p)}\right)$$

$$P(a < X \leq b) \approx F_X(b) - F_X(a) \quad \text{(för normalfördelning)}$$
$$= \Phi\left(\frac{b-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a-np}{\sqrt{np(1-p)}}\right)$$

Tilläggsinformation

Exempel för normalapproximation: Massproduktion av byggelement. Ett byggelement blir defekt med $p = 0.1$. En byggmästare köper 1000 stycken.

Sannolikheten att han får minst 80, högst 120 felaktiga ?

Slumpvariabel X: antalet felaktiga han köpt $\Omega = \{0, 1, \ldots 1000\}$

$$X \in Bin(n,p) = Bin(1000, 0.1) \quad \text{med} \quad np(1-p) = 90 > 10$$
$$X \in N\left(n \cdot p, \sqrt{np(1-p)}\right) = N\left(1000 \cdot 0.1, \sqrt{90}\right) = N(100, 9.487)$$

$$\begin{aligned} P(\mathbf{79} < X \leq 120) &= F_X(120) - F_X(79) \\ &= \Phi\left(\frac{120-100}{9.487}\right) - \Phi\left(\frac{79-100}{9.487}\right) \\ &= \Phi(2.11) - \Phi(-2.21) \\ &= 0.9826 + 0.98645 - 1 = 0.969 \end{aligned}$$

Tilläggsinformation: halvkorrektion

Istället för:

$$P(a < X \leq b) \approx \Phi\left(\frac{b-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a-np}{\sqrt{np(1-p)}}\right)$$

kan man räkna:

$$P(a < X \leq b) \approx \Phi\left(\frac{b+\mathbf{1/2}-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a+\mathbf{1/2}-np}{\sqrt{np(1-p)}}\right)$$

för att det blir mera noggrannt ..