Binomialverteilung, Konfidenzintervall

(Würfeln, sechs oder nicht sechs)

Prof. Dr. Dörte Haftendorn 9.5.08 MuPAD 4 Update von 2011
http:haftendorn.uni.leuphana.de www.mathematik-verstehen.de

++

3. Konfidenzintervalle

4. Betrachtung der Verteilung

###

3. Konfidenzintervall 5%-Niveau

```
fn := stats::normalQuantile(0, 1):
z1:=fn(0.995);
z5:=fn(0.975)
```

$$2.575829304$$

$$1.959963985$$

Konfidenzintervall-Ansatz

```
confAnsatz:=(k/n-p)^2=z5*p*(1-p)/n
```

$$\left(p - \frac{k}{n}\right)^2 = -\frac{1.959963985 \cdot p \cdot (p-1)}{n}$$

```
gl:=confAnsatz|{n=300,k=46.0}
```

$$(p - 0.1533333333)^2 = -0.006533213282 \cdot p \cdot (p-1)$$

```
solve(gl,p)
```

$$\{0.1264679479,\ 0.1846990119\}$$

Antwort: Aus dem Versuch mit n=300 erhält man ein 5%-Konfidenzintervall von 12,6% < p<18.5%

```
gl:=confAnsatz|{n=3000,k=527}
```

$$\left(p - \frac{527}{3000}\right)^2 = -0.0006533213282 \cdot p \cdot (p-1)$$

```
solve(gl,p)
```

$$\{0.1661527265,\ 0.1856041179\}$$

Antwort: Aus dem Versuch mit n=3000 erhält man ein 5%-Konfidenzintervall von 16,6% < p<18,5%

```
gl:=confAnsatz|{n=30000,k=5065}
```

$$\left(p - \frac{1013}{6000}\right)^2 = -0.00006533213282 \cdot p \cdot (p-1)$$


```
solve(gl,p)
```

$$\{0.165827124,\ 0.1718828115\}$$

$\{0.165827124, 0.1718828115\}$

Antwort: Aus dem Versuch mit n=30000 erhält man ein 5%-Konfidenzintervall von 16,58% < p<17,19%

Spielplatz

```
gl:=confAnsatz|{n=3000,k=500};
solve(gl,p)
```

$$\left(p - \frac{1}{6}\right)^2 = -0.0006533213282 \cdot p \cdot (p-1)$$

$\{0.1573592161, 0.1764093804\}$

Andersherum, direkte Deutung der Binomialverteilung:

```
gl:=confAnsatz|{n=3000,k=500};
solve(gl,p)
```

$$\left(p - \frac{1}{6}\right)^2 = -0.0006533213282 \cdot p \cdot (p-1)$$

$\{0.1573592161, 0.1764093804\}$

Bei 3000 Wurf liegen 95% solcher Simulationen in dem genannten Bereich.

##

##

4. Betrachtung der Verteilung und Quantile

```
nn:=300: pp:=1/6.0; my:=nn*pp;sig:=sqrt(nn*pp*(1-pp));
xmin:=my-4*sig;xmax:=my+4*sig;
```

0.1666666667

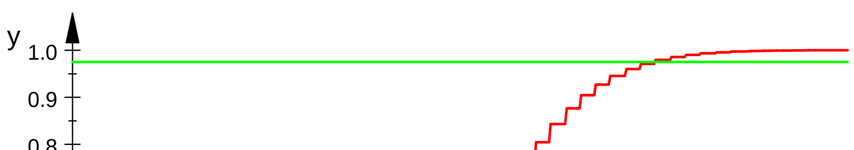
50.0

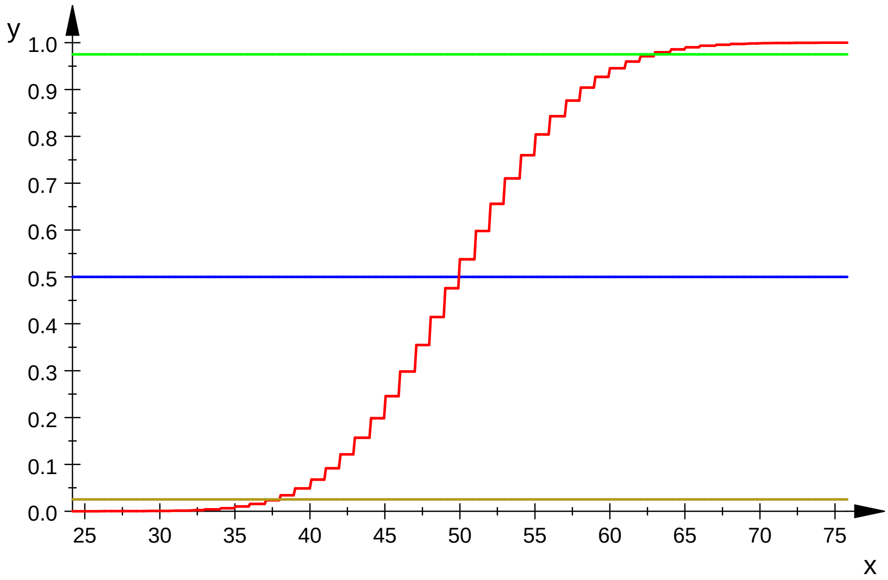
6.454972244

24.18011103

75.81988897

```
bicdf:=stats::binomialCDF(nn,pp):
plotfunc2d(0.5,bicdf(x),0.975,0.025,x=xmin..xmax,
  LegendVisible=FALSE)
```

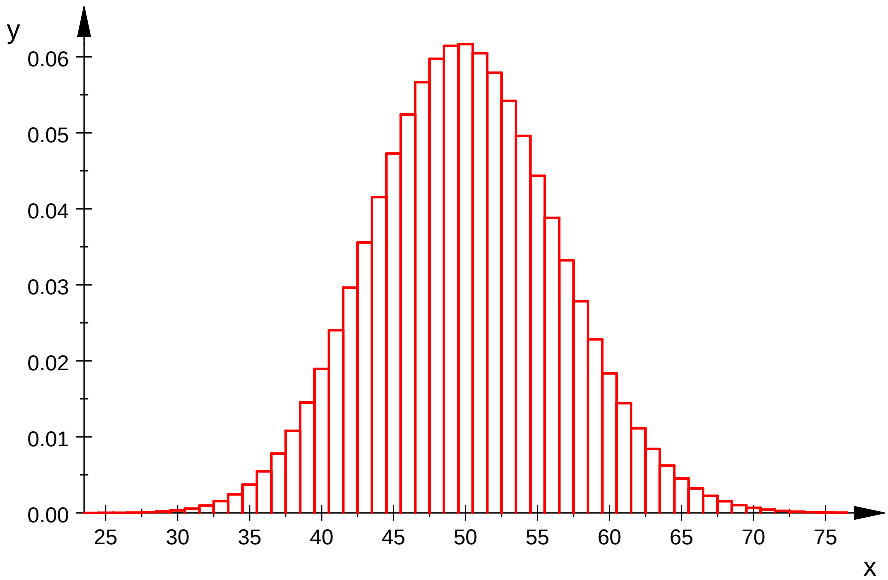





biHist ist eine Zeichenprozedur die bei Datei-Eigenschaften eingetragen ist un daher hier aufgerufen werden kann

```
w=1 alle Werte, w=0 nur my, sigma
```

```
biHist(nn,pp,xmin,xmax,0)
```



["my", 50.0], ["sigma", 6.454972244]

```
nn:=3000: pp:=1/6.0; my:=nn*pp;sig:=sqrt(nn*pp*(1-pp));
xmin:=my-4*sig;xmax:=my+4*sig;
```

0.1666666667

500.0

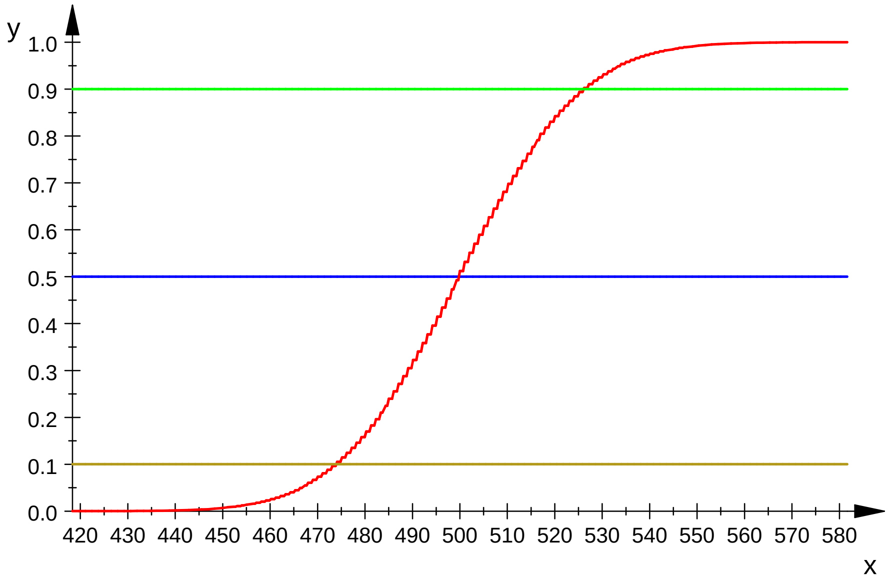
20.41241452

418.3503419

581.6496581

581.6496581

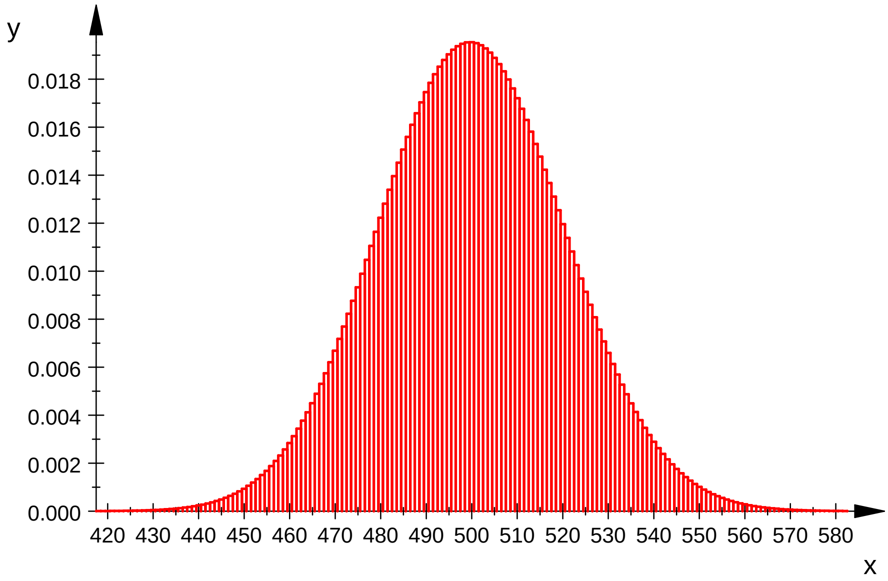
```
bicdf:=stats::binomialCDF(nn,pp):
plotfunc2d(0.5,bicdf(x),0.9,0.1,x=xmin..xmax,
  LegendVisible=FALSE)
```



biHist ist eine Zeichenprozedur die bei Datei-Eigenschaften eingetragen ist un daher hier aufgerufen werden kann

```
w=1 alle Werte, w=0 nur my, sigma
```

```
biHist(nn,pp,xmin,xmax,0)
```



["my", 500.0], ["sigma", 20.41241452]

```
nn:=30000: pp:=1/6.0; my:=nn*pp;sig:=sqrt(nn*pp*(1-pp));
xmin:=my-4*sig;xmax:=my+4*sig;
```

0.1666666667

5000.0

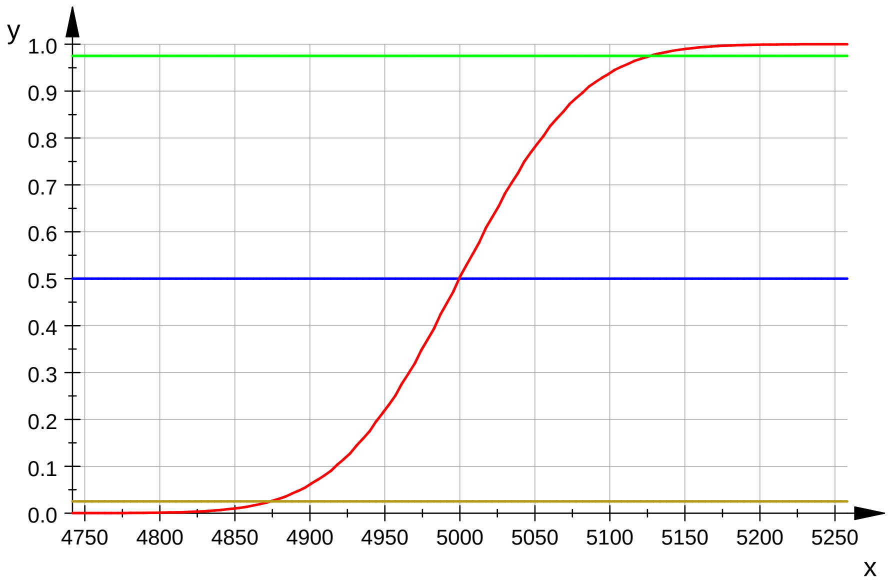
5000.0

64.54972244

4741.80111

5258.19889

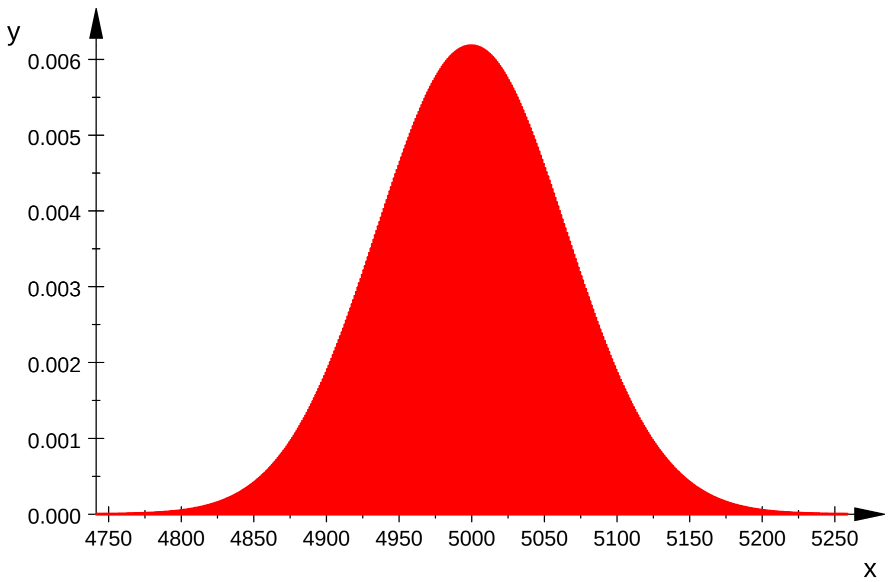
```
bicdf:=stats::binomialCDF(nn,pp):
plotfunc2d(0.5,bicdf(x),0.975,0.025,x=xmin..xmax,
  LegendVisible=FALSE, GridVisible=TRUE)
```



biHist ist eine Zeichenprozedur die bei Datei-Eigenschaften eingetragen ist und daher hier aufgerufen werden kann

```
w=1 alle Werte, w=0 nur my, sigma
```

```
biHist(nn,pp,xmin,xmax,0)
```



["my", 5000.0], ["sigma", 64.54972244]

Betrachtung der Quantile

Betrachtung der Quantile

Angabe der Intervalle, in denen 95% solcher Versuche
beim exakten Würfel wohl liegen werden
absolut und relativ

```
nn:=300: pp:=1/6.0:kf:=0.95:
f := stats::binomialQuantile(nn, pp):
grenzen:=f((1-kf)/2),f((1+kf)/2);
float(grenzen[1]/nn),float(grenzen[2]/nn);
```

$$38, 63$$

$$0.1266666667, 0.21$$

```
nn:=3000: pp:=1/6.0:kf:=0.95:
f := stats::binomialQuantile(nn, pp):
grenzen:=f((1-kf)/2),f((1+kf)/2);
float(grenzen[1]/nn),float(grenzen[2]/nn);
```

$$460, 540$$

$$0.1533333333, 0.18$$

Bestimmung eines 95%-Intervalls aus der Normalverteilung

```
z5;
gl:=confAnsatz|{n=3000,k=500};
solve(gl,p)
```

$$1.959963985$$

$$\left(p-\frac{1}{6}\right)^2 = -0.0006533213282 \cdot p \cdot (p-1)$$

$$\{0.1573592161, 0.1764093804\}$$

Bei 3000 Wurf liegen 95% solcher Simulationen in dem genannten Bereich.
Die Methoder mit den Binomialquantilen ist aber genauer.
##

```
nn:=30000: pp:=1/6.0:kf:=0.95:
f := stats::binomialQuantile(nn, pp):
grenzen:=f((1-kf)/2),f((1+kf)/2);
float(grenzen[1]/nn),float(grenzen[2]/nn);
```

$$4874, 5127$$

$$0.1624666667, 0.1709$$

Das heißt, dass man bei im Mittel bei einem von 20 solchen Versuchen
außerhalb des genannten Bereichs landen wird.