Vorlesung vom 31.07.2013:

- Hypergeometrische Verteilung
- Urnenmodelle (im Rückblick)
- Noch einmal Binomialverteilung

Hypergeometrische Verteilung: Thema Stichprobenauswahl im Fall von „2 Sorten" von Objekten z.B.:

Ω: Stichprobenraum als Gruppe von Personen, z.B. Fallunterscheidung „studiert Biologie" vs. „studiert nicht Biologie" „Dualität"

$B \subseteq \Omega$: Ereignis, dass $\omega \in \Omega$ Biologie studiert

$B^c = \Omega \setminus B$: Komplementärereignis, dass $\omega \in \Omega$ nicht Biologie studiert.

Falls $|\Omega| = N$ Gesamtanzahl an Personen, greift man eine Stichprobe der Größe $n < N$ zufällig heraus.

Falls $|B| = r < N$, kann man nach der Wahrscheinlichkeit fragen, dass innerhalb der Stichprobe x Personen zu B gehören.

„Wir" nimmt häufig eine Zufallsvariable $X: \Omega_n \to R$ für

$\Omega_n = \{E_n = \underbrace{\{\omega_1, \dots, \omega_n\}}_{\text{gezogene Stichprobe}} \mid \omega_i \in \Omega\}$ als Menge aller n-elementigen Teilmengen (=Stichproben) in Ω.

$X(E_n)=x$, Anzahl der $w_i \in E_n$ mit $w_i \in B$

Gesucht: Wahrscheinlichkeit $\boxed{P(X=x)=P(\{E_n\in\Omega_n \mid X(E_n)=x\})}$

zufällige Stichprobenwahl = Laplace-Experiment!

Frage: Welches Urnenmodell liegt dieser Stichprobenwahl zugrunde?

Antwort: Ziehen ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge $\Rightarrow$ Binomialkoeffizienten nötig!!

Variablen / Parameter:

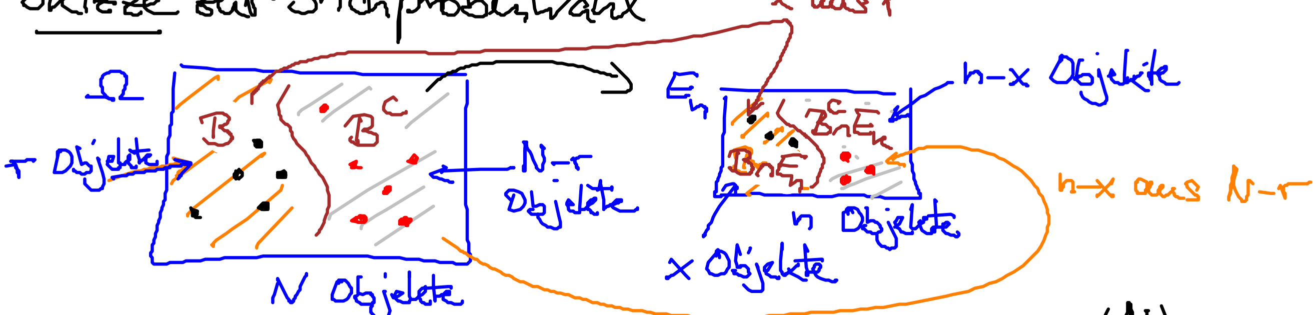
$N=|\Omega|$: Anzahl Gesamtmenge, aus der die Stichprobe gezogen wird

$r=|B|$: Anzahl an $w\in\Omega$, welche das „Merkmal" B haben, d.h. für die $w\in B$ gilt.

$n=|E_n|$: Stichprobengröße

x : Anzahl der $w_i\in E_n$ mit $w_i\in B$

Skizze zur Stichprobenwahl



Gesamtanzahl an n-elementigen Stichproben: $|\Omega_n|=\binom{N}{n}$

„Günstige" Ergebnisse: Stichproben $E_n\in\Omega_n$, welche genau x Objekte $w_i\in B$ enthält. Wieviele sind derart möglich?

$$\boxed{P(X=x)=\frac{\binom{r}{x}\cdot\binom{N-r}{n-x}}{\binom{N}{n}}}$$

hypergeometrische Verteilung!!
Vorgegeben ist hier: N, n, r

Jetzt Anwendung:

Ü 57 Parameter: $N=10$ $(=6+4)$, Ω: Menge der 10 Früchte (Obstkiste)

$n=5$ (Stichprobengröße), $r=6$ für „Merkmal" A: „ein Apfel"

hier: $0\le x\le 2$, d.h. $x=0,1,2$

Zufallsgröße: $X: \Omega_5 \to \mathbb{R}$ mit $X(E_5) = x$ zählt die Anzahl an $\omega_i \in A$ in $E_5 \in \Omega_5$ ($n=5$)

Rechnung: $\boxed{P(X \leq 2)} = P(X=0) + P(X=1) + P(X=2)$

$\{X=0\}$ und $\{X=1\}$ und $\{X=2\}$ sind disjunkte Ereignisse

$$= \frac{\binom{6}{0}\binom{4}{5}}{\binom{10}{5}} + \frac{\binom{6}{1}\binom{4}{4}}{\binom{10}{5}} + \frac{\binom{6}{2}\binom{4}{3}}{\binom{10}{5}}$$

geht doch gar nicht!

Beachte beim Binomialkoeffizienten: $\binom{4}{5} = \frac{4 \cdot 3 \cdot 2 \cdot 1 \cdot 0}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 0$

allgemein: $\binom{\alpha}{k} = \frac{\alpha \cdot (\alpha-1) \cdot \ldots \cdot (\alpha-k+1)}{\underbrace{1 \cdot 2 \cdot 3 \cdot \ldots \cdot k}_{k \text{ Faktoren}}}$

$$\binom{10}{5} = \frac{\cancel{10} \cdot \cancel{9} \cdot \cancel{8} \cdot 7 \cdot 6}{1 \cdot \cancel{2} \cdot \cancel{3} \cdot \cancel{4} \cdot \cancel{5}} = 6 \cdot 7 \cdot 6 = 7 \cdot 36 = 252$$

$$= \frac{1 \cdot 0}{252} + \frac{6 \cdot 1}{252} + \frac{15 \cdot 4}{252}$$

$$= 0 + \frac{6+60}{252} = \frac{66}{252} = \frac{264}{1008} \approx 0{,}262 \mathrel{\hat{=}} 26{,}2\%$$

$\cdot \frac{4}{4}$

$\binom{4}{4} = \binom{4}{0} = 1$, $\binom{4}{3} = \binom{4}{1} = 4$

$\binom{n}{k} = \binom{n}{n-k}$

$\binom{6}{2} = \frac{\cancel{6}^{3} \cdot 5}{1 \cdot \cancel{2}} = 15$

Fall: Ziehung ohne Zurücklegen!!

Jetzt: Ziehung mit Zurücklegen (Fall (b)): Bernoulli-Experiment $\Rightarrow$ Binomialverteilung!!

$n=5$ Ziehungen, gefragt: Anzahl $X=x$ an „Erfolgen" (=Apfel)

Wir benötigen noch p als Wahrscheinlichkeit für „Erfolg", $q = 1-p$ als Wahrscheinlichkeit für „Misserfolg".

Hier: $p = \frac{6}{10} = 0{,}6$, $q = \frac{4}{10} = 0{,}4 = 1-p$.

Allgemein Bino-

...mialverteilung: $P(X=x) = \binom{n}{x} \cdot p^x q^{n-x}$

Dann: $\boxed{P(X \leq 2)} = P(X=0) + P(X=1) + P(X=2)$

$$= \binom{5}{0} \cdot 0{,}6^0 \cdot 0{,}4^5 + \binom{5}{1} \cdot 0{,}6 \cdot 0{,}4^4 + \binom{5}{2} 0{,}6^2 \cdot 0{,}4^3$$

$$= 1 \cdot 1 \cdot 0{,}4^5 + 5 \cdot 0{,}6 \cdot 0{,}4^4 + 10 \cdot \underbrace{0{,}6^2}_{0{,}36} \cdot 0{,}4^3$$

$$= (\underbrace{0{,}4^2}_{0{,}16} + \underbrace{3{,}0 \cdot 0{,}4}_{1{,}2} + 3{,}6) \cdot 0{,}4^3 = \overbrace{4{,}96 \cdot 0{,}064}^{=0{,}317\ldots}$$

$$\simeq 5 \cdot 0{,}06 = 0{,}3 \mathrel{\hat=} 30\,\%$$

Fazit: Der Unterschied zwischen hypergeometrischer und Binomialverteilung ist nicht allzu groß, wenn die Stichprobengröße n im Verhältnis zur Gesamtanzahl N relativ klein ist, denn die Wahrscheinlichkeit bei geringer Anzahl von Ziehungen gegenüber der Anzahl $N = |\Omega|$ bei Zurücklegen ein Objekt $\omega \in \Omega$ doppelt zu ziehen, ist ja fast $= 0$.

!!

Daher wird bei kleiner Stichprobengröße und $N = |\Omega|$ groß gern die Binomial- anstelle der hypergeometrischen Wahrscheinlichkeit genommen.

Noch einmal beide Rechnungen in der Aufgabe, wenn eine Birne durch einen Apfel ersetzt wird:

(a) hypergeometrische V. : $N = 10,\ n = 5,\ r = 7$

(= Ziehen ohne Zurücklegen)

$$\frac{7 \cdot \cancel{6}^{\,3}}{1 \cdot \cancel{2}_{\,1}} = 21$$

$$\boxed{P(X \leq 2)} = \frac{1}{\underbrace{\binom{10}{5}}_{252}} \cdot \left\{ \binom{7}{0} \cdot \underbrace{\binom{3}{5}}_{0} + \binom{7}{1} \cdot \underbrace{\binom{3}{4}}_{0} + \underbrace{\binom{7}{2}}_{21} \cdot \underbrace{\binom{3}{3}}_{1} \right\} = \frac{21}{252}$$

$$= 0{,}0833\ldots \simeq 0{,}083 \mathrel{\hat=} 8{,}3\,\%$$

(b) Binomialvert. :

(= Ziehen mit Zurücklegen)

$n = 5,\ p = \frac{7}{10} = 0{,}7;\ q = \frac{3}{10} = 0{,}3$

$$\boxed{P(X \leq 2)} = \underbrace{\binom{5}{0}}_{1} \cdot \underbrace{0{,}7^0}_{1} \cdot 0{,}3^5 + \underbrace{\binom{5}{1}}_{5} 0{,}7^1 \cdot 0{,}3^4 + \underbrace{\binom{5}{2}}_{10} \cdot 0{,}7^2 \cdot 0{,}3^3$$

$$= (\underbrace{1 \cdot 0{,}3^2}_{0{,}09} + \underbrace{5 \cdot 0{,}7 \cdot 0{,}3}_{0{,}21} + 10 \cdot \underbrace{0{,}7^2}_{0{,}49}) \cdot 0{,}3^3 = (0{,}09 + 1{,}05 + 4{,}9) \cdot \underbrace{0{,}3^3}_{0{,}027}$$

$$= 6{,}04 \cdot 0{,}027 = 0{,}16308$$

$$\simeq 0{,}163 \mathrel{\hat=} 16{,}3\,\%$$

ENDE der Vorlesung !