

<b>Ue Statistik u. Wahrscheinlichkeitsth. f. Inf.</b> 107.251      W 2005 6 <a href="http://www.statistik.tuwien.ac.at/RISueb/">http://www.statistik.tuwien.ac.at/RISueb/</a>	Di 12:00 – 18:00
	<b>3.Blatt</b>
Werner GURKER      Tel.: 58801-10724 Spr.: Di u. Do 11-12      e-mail: W.Gurker@tuwien.ac.at	25. Okt. 2005

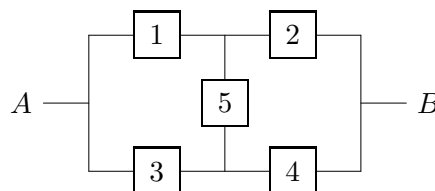
- Zwei Zahlen  $a$  und  $b$  werden zufällig (mit Zurücklegen) aus der Menge  $\{1, 2, 3, \dots, k\}$  ausgewählt.
  - Wenn  $k = 20$ , mit welcher (exakten) Wahrscheinlichkeit sind  $a$  und  $b$  relativ prim (teilerfremd)?
  - \*Ermitteln Sie einen approximativen Ausdruck (nicht berechnen) für diese Wahrscheinlichkeit, wenn  $k$  eine große (natürliche) Zahl ist. Was ergibt sich für  $k \rightarrow \infty$ ?
- Bei einer Lotterie mit 40000 Losen gibt es drei große Gewinne.
  - Bestimmen Sie die Wahrscheinlichkeit, mit der man zumindest einen großen Gewinn pro 1000 Losen bekommt.
  - \*Wie viele Lose müßte man kaufen, damit die Wahrscheinlichkeit eines großen Gewinns zumindest 0.5 ist?
- (a) Zeigen Sie die *Boole'sche Ungleichung*:

$$W\left(\bigcup_{i=1}^n E_i\right) \leq \sum_{i=1}^n W(E_i)$$

- (b) Zeigen Sie die *Bonferroni'sche Ungleichung*:

$$W\left(\bigcap_{i=1}^n E_i\right) \geq 1 - \sum_{i=1}^n W(\overline{E}_i)$$

- In der folgenden Brückenschaltung sind die Verbindungen (unabhängig voneinander) mit gleichbleibender Wahrscheinlichkeit  $p$  ( $0 \leq p \leq 1$ ) intakt. Die Schaltung ist intakt, wenn Strom von  $A$  nach  $B$  fließen kann.



$E$  sei das Ereignis, daß die Schaltung intakt ist, und  $V_i$  das Ereignis, daß Verbindung  $i$  intakt ist ( $i = 1, \dots, 5$ ).

- Wie lautet ein passender Merkmalraum? Wie läßt sich  $V_i$  auf Basis dieses Merkmalraums ausdrücken?
- Drücken Sie  $E$  mit Hilfe von  $V_i$ ,  $i = 1, \dots, 5$ , aus.

- (c) Benützen Sie die Darstellung von (b) um  $W(E)$  mit Hilfe des Additionstheorems zu berechnen.
  - \*(d) Berechnen Sie  $W(E)$  mit Hilfe des Satzes von der vollständigen Wahrscheinlichkeit, indem Sie durch den Status (intakt/defekt) von Verbindung 5 bedingen.
5. Einer von zwei Behältern mit jeweils 10 Kugeln enthält eine markierte Kugel. Es ist bekannt, daß sich die markierte Kugel mit Wahrscheinlichkeit  $2/3$  im ersten Behälter befindet. Sie bekommen die Möglichkeit, hintereinander 20 Kugeln mit Zurücklegen nach Belieben aus den Behältern zu ziehen. Wie gehen Sie vor, wenn Sie die markierte Kugel mit möglichst hoher Wahrscheinlichkeit finden möchten?
- (a) Angenommen, Sie entscheiden sich dafür, 10 Kugeln aus dem ersten und 10 Kugeln aus dem zweiten Behälter zu ziehen. Mit welcher Wahrscheinlichkeit finden Sie die markierte Kugel?
  - \*(b) Wie lautet die optimale Aufteilung der Ziehungen?
6. Angenommen, von 100 VLSI-Chips sind im Durchschnitt 3 defekt. Ein automatischer Test erkennt einen guten Chip in 90% der Fälle, und einen defekten Chip in 95% der Fälle als solchen.
- (a) Wie groß ist die Wahrscheinlichkeit, daß ein als gut erkannter Chip tatsächlich gut ist?
  - (b) Wie groß ist die Wahrscheinlichkeit, daß ein als defekt erkannter Chip tatsächlich defekt ist? Geben Sie eine Erklärung für die (unerwartet?) kleine Wahrscheinlichkeit.

Beispiel(teil)e mit (\*) werden im Konversatorium behandelt.

1. (a) Der Merkmalraum des Versuchs ist  $M = \{(x, y) : x, y = 1, 2, \dots, 20\}$ ; er besteht aus  $m = 20^2 = 400$  Elementen. Zwei Zahlen sind relativ prim, wenn sie keine gemeinsamen Teiler haben; dabei kann man sich auf die Primzahlen kleiner gleich 20 beschränken. Die Zahl der günstigen Versuchsausgänge läßt sich durch einfaches Abzählen oder – etwas schneller – mit Hilfe des In- und Exklusionsgesetzes (vgl. Lösung zu **Bsp 1.4**, Bemerkung) bestimmen. Bezeichnet man mit  $E_p$  das Ereignis, daß beide Zahlen durch  $p$  teilbar sind, so gilt mit  $I = \{2, 3, 5, 7, 11, 13, 17, 19\}$ :

$$\begin{aligned} \left| \bigcup_{p \in I} E_p \right| &= \sum_{p \in I} |E_p| - |E_2 \cap E_3| - |E_2 \cap E_5| - |E_2 \cap E_7| - |E_3 \cap E_5| \\ &= 100 + 36 + 16 + 4 + 1 + 1 + 1 + 1 - 9 - 4 - 1 - 1 \\ &= 145 \end{aligned}$$

Man beachte, daß alle nicht hingeschriebenen paarweisen Durchschnitte und alle Durchschnitte höherer Ordnung leer sind. Die gesuchte Wahrscheinlichkeit ist also gegeben durch:

$$\frac{g}{m} = \frac{400 - 145}{400} = \frac{255}{400} = \frac{51}{80} = 0.6375$$

- \*(b) Ist  $p_1, p_2, \dots$  die Folge der Primzahlen, so ist ein approximativer Ausdruck für die gesuchte Wahrscheinlichkeit für großes  $k$  gegeben durch:

$$\prod_{p_i \leq k} \left[ 1 - \left( \frac{k/p_i}{k} \right)^2 \right] = \prod_{p_i \leq k} \left( 1 - \frac{1}{p_i^2} \right)$$

Für  $k \rightarrow \infty$  gilt:

$$\lim_{k \rightarrow \infty} \prod_{p_i \leq k} \left( 1 - \frac{1}{p_i^2} \right) = \prod_{i=1}^{\infty} \left( 1 - \frac{1}{p_i^2} \right) = \frac{6}{\pi^2} \doteq 0.6079$$

Bem.: Allgemeiner gilt:

$$\prod_{i=1}^{\infty} \left( 1 - \frac{1}{p_i^s} \right) = \frac{1}{\zeta(s)}$$

Dabei ist  $\zeta(s)$  die *Riemann'sche Zeta-Funktion*:

$$\zeta(s) = \sum_{i=0}^{\infty} \frac{1}{i^s}$$

Sie konvergiert für  $s > 1$ ; für  $s = 2$  gilt:

$$\zeta(2) = \sum_{i=0}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6}$$

2. (a) Gibt es  $N$  ( $= 40000$ ) Lose und  $k$  ( $= 3$ ) große Gewinne, so ist die Wahrscheinlichkeit, bei  $n$  ( $= 1000$ ) Ziehungen zumindest einen großen Gewinn zu bekommen, gegeben durch:

$$p = 1 - \frac{N-k}{N} \cdot \frac{N-k-1}{N-1} \cdots \frac{N-k-n+1}{N-n+1} = 1 - \frac{(N-k)!(N-n)!}{N!(N-k-n)!}$$

Da  $k$  im Vergleich zu  $N$  sehr klein ist, läßt sich dieser Ausdruck approximativ auch wie folgt schreiben:

$$p \approx 1 - \left(\frac{N-n}{N}\right)^k$$

Konkret ergibt sich:

$$p \approx 1 - \left(\frac{39}{40}\right)^3 = 0.0731$$

\*(b) Dafür muß gelten:

$$\frac{1}{2} \geq \left(\frac{N-n}{N}\right)^k \implies n \geq N \left[1 - \left(\frac{1}{2}\right)^{1/k}\right]$$

Konkret ergibt sich:

$$n \geq 40000 \left[1 - \left(\frac{1}{2}\right)^{1/3}\right] \doteq 8252$$

3. (a) Die Vereinigung der Mengen  $E_i$ ,  $i = 1, \dots, n$ , läßt sich als Vereinigung von (paarweise) disjunkten Mengen darstellen:

$$\bigcup_{i=1}^n E_i = \bigcup_{i=1}^n \left( E_i \cap \bigcap_{j=1}^{i-1} \overline{E}_j \right)$$

Damit folgt die Behauptung:

$$W\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n W\left(E_i \cap \bigcap_{j=1}^{i-1} \overline{E}_j\right) \leq \sum_{i=1}^n W(E_i)$$

Die Ungleichung gilt wegen  $A \subseteq B \implies W(A) \leq W(B)$ .

(b) Nach den DeMorgan'schen Regeln gilt:

$$\bigcap_{i=1}^n E_i = \overline{\left(\bigcup_{i=1}^n \overline{E}_i\right)}$$

Daraus folgt mit  $W(\overline{A}) = 1 - W(A)$  und (a) die Behauptung:

$$W\left(\bigcap_{i=1}^n E_i\right) = 1 - W\left(\bigcup_{i=1}^n \overline{E}_i\right) \geq 1 - \sum_{i=1}^n W(\overline{E}_i)$$

4. (a) Kodiert man mit 0/1 eine defekte/intakte Verbindung, so ist ein geeigneter Merkmalraum gegeben durch:

$$M = \{(x_1, x_2, x_3, x_4, x_5) \mid x_i = 0, 1; i = 1, \dots, 5\}$$

Für das Ereignis  $V_i$  (= „Verbindung  $i$  ist intakt“) gilt dann:

$$V_i = \{(x_1, x_2, x_3, x_4, x_5) \mid x_i = 1, x_j = 0, 1, j \neq i\} \subseteq M$$

- (b) Das (zusammengesetzte) Ereignis  $E$  (= „Strom fließt von  $A$  nach  $B$ “) lässt sich auf Basis der Ereignisse  $V_i$  wie folgt ausdrücken:

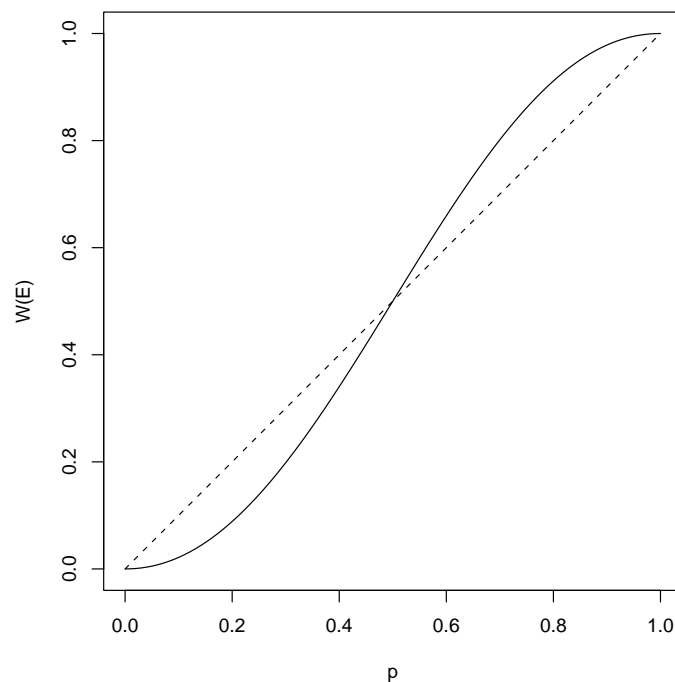
$$E = (V_1 \cap V_2) \cup (V_3 \cap V_4) \cup (V_1 \cap V_4 \cap V_5) \cup (V_2 \cap V_3 \cap V_5)$$

- (c) Mit dem Additionstheorem folgt (die Ereignisse  $V_i$  sind nach Voraussetzung unabhängig):

$$W(E) = 2p^2 + 2p^3 - 5p^4 - p^5 + 4p^5 - p^5 = 2p^2 + 2p^3 - 5p^4 + 2p^5$$

- \*(d) Bedingen durch den Status von Verbindung 5 (Satz von der vollständigen Wahrscheinlichkeit):

$$\begin{aligned} W(E) &= W(E|V_5)W(V_5) + W(E|\overline{V_5})W(\overline{V_5}) \\ &= W((V_1 \cup V_3) \cap (V_2 \cup V_4))W(V_5) + W((V_1 \cap V_2) \cup (V_3 \cap V_4))W(\overline{V_5}) \\ &= [2p - p^2]^2 p + [2p^2 - p^4](1 - p) \\ &= 2p^2 + 2p^3 - 5p^4 + 2p^5 \end{aligned}$$



5. (a) Bezeichnet  $H_1$  das Ereignis, daß die markierte Kugel im ersten Behälter ist, und  $H_2$  das Ereignis, daß sie im zweiten ist, und ist  $A$  das Ereignis, daß sie gefunden wird, so gilt mit dem Satz von der vollständigen Wahrscheinlichkeit:

$$\begin{aligned} W(A) &= W(A|H_1)W(H_1) + W(A|H_2)W(H_2) \\ &= \left[1 - \left(\frac{9}{10}\right)^{10}\right] \frac{2}{3} + \left[1 - \left(\frac{9}{10}\right)^{10}\right] \frac{1}{3} \\ &= 1 - \left(\frac{9}{10}\right)^{10} = 0.6513 \end{aligned}$$

- \*(b) Bezeichnen wir allgemeiner mit  $p$  die Wahrscheinlichkeit, daß die markierte Kugel im ersten Behälter ist, und machen wir insgesamt  $n$  Ziehungen mit Zurücklegen ( $m$  aus dem ersten Behälter,  $n - m$  aus dem zweiten), so gilt analog zu (a):

$$W(A) = \left[1 - \left(\frac{9}{10}\right)^m\right] p + \left[1 - \left(\frac{9}{10}\right)^{n-m}\right] (1 - p)$$

Für konkrete Werte von  $n$  und  $p$  läßt sich das optimale  $m$  durch Probieren finden oder durch die folgende Überlegung ( $da^x/dx = a^x \ln a$ ):

$$\frac{dW(A)}{dm} = -p \left(\frac{9}{10}\right)^m \ln \left(\frac{9}{10}\right) + (1 - p) \left(\frac{9}{10}\right)^{n-m} \ln \left(\frac{9}{10}\right)$$

Setzt man die Ableitung gleich Null, so ergibt sich:

$$\left(\frac{9}{10}\right)^{2m-n} = \frac{1-p}{p} \implies m = \frac{n}{2} + \frac{\ln[(1-p)/p]}{2 \ln(9/10)}$$

Im vorliegenden Fall ergibt sich  $m = 13.23$ ; die optimale Aufteilung ist also 13 aus dem ersten Behälter und 7 aus dem zweiten. Die Wahrscheinlichkeit, daß die markierte Kugel gefunden wird, beträgt dann 0.6711.

6. Bezeichnet man mit  $C$  das Ereignis, daß ein Chip gut ist, und mit  $G$  das Ereignis, daß der Test den Chip als gut ausweist, so gilt laut Angabe:

$$W(C) = 0.97, \quad W(G|C) = 0.9, \quad W(\overline{G}|\overline{C}) = 0.95$$

Die Wahrscheinlichkeiten bestimmt man nun mittels Bayes'scher Formel:

(a)

$$\begin{aligned} W(C|G) &= \frac{W(G|C)W(C)}{W(G|C)W(C) + W(G|\overline{C})W(\overline{C})} \\ &= \frac{0.9 \cdot 0.97}{0.9 \cdot 0.97 + 0.05 \cdot 0.03} = 0.9983 \end{aligned}$$

(b)

$$W(\overline{C}|\overline{G}) = \frac{0.95 \cdot 0.03}{0.95 \cdot 0.03 + 0.1 \cdot 0.97} = 0.2271$$