Das Blutuntersuchungsproblem

Aufgabenstellung:

Durch eine Blutuntersuchung soll bei einer großen Anzahl von Personen geprüft werden, ob jemand an einer bestimmten Krankheit leidet oder nicht.

Zum Nachweis des Krankheitserregers steht ein Test zur Verfügung.

Mit der Wahrscheinlichkeit p ist eine Person gesund, mit der Wahrscheinlichkeit (1 - p) krank.

Testverfahren:

a) *Einzelprüfung*: Man benötigt für n Personen n Tests, d.h. für jede Person einen Test.

b) *Gruppenprüfung*: Das Blut von n Personen wird vermischt. Wird der Erreger nicht nachgewiesen, so hat man für diese n Personen nur 1 Test benötigt, wird der Erreger jedoch im Mischblut nachgewiesen, so muss man zur Einzelprüfung übergehen und hat dann insgesamt für n Personen (n+1) Test benötigt.

Für den Erwartungswert μ = E(X) der notwendigen Blutanalysen im Fall der Gruppenprüfung für eine Gruppenanzahl der Größe n gilt:

$$\begin{aligned} \boldsymbol{\mu} &= 1 \cdot p^n + (n+1) \cdot (1-p^n) \\ &= n+1 - n \cdot p^n \end{aligned}$$

Da man bei Einzelprüfung genau n Analysen benötigt, gilt für die absolute Ersparnis S(n):

$$S(n) := n - \boldsymbol{\mu} = n \cdot p^n - 1$$

Betrachtet man die relative Ersparnis **s**(n), d.h. die Ersparnis pro Person, so gilt:

$$\mathbf{s}(n) = p^n - \frac{1}{n}$$

Für welche Gruppengröße n ist **s** maximal ?

Trivialerweise muß gelten:

$$p > \frac{1}{\sqrt[n]{n}} \quad \left[= n^{-\frac{1}{n}} = e^{\frac{-\ln(n)}{n}} \right]$$

n	1	2	3	4	5	6	7	8	9	10	11
$\frac{1}{\sqrt[n]{n}}$	1,000	0,707	0,693	0,707	0,725	0,742	0,757	0,771	0,783	0,794	0,804

Das Blutuntersuchungsproblem

Grundsätzlich hat man nur eine Ersparnis, wenn $p > 0{,}693$ ist !

Für z.B. p = 70 % ist die optimale Gruppengröße 3, wobei die Ersparnis gering ist. (Die Zweiergruppe ist niemals optimal!)

Es sei p = 99%:[1]

n	3	4	5	6	7	8	9	10	11	12	13
s(n)	0,637	0,711	0,751	0,775	0,789	0,798	0,802	0,804	0,804	0,803	0,801

Lösung mit Differentialrechnung:

$$\mathbf{s}'(n) = \ln(p) \cdot p^n + \frac{1}{n^2} = 0 \quad \Leftrightarrow \quad n = \sqrt{\frac{-1}{\ln(p) \cdot p^n}}$$

k	n_k	$\varphi(n_k)$
1	9,00000000	10,43641492
2	10,43641492	10,51201986
3	10,51201986	10,51601442
4	10,51601442	10,51622552
5	10,51622552	10,51623667

[1] Der reale Hintergrund dieser Aufgabe ist die ärztliche Untersuchung der amerikanischen Rekruten im 2.Weltkrieg auf Syphilis. Die Wahrscheinlichkeit, an dieser Krankheit zu leiden, betrug seinerzeit ungefähr 1%.

Quelle: Arthur Engel: Wahrscheinlichkeitsrechnung und Statistik, Band 1; Verlag: Klett 1973