

Kansrekening en Statistiek

Henk Broer

Instituut voor Wiskunde en Informatica
Rijksuniversiteit Groningen

Overzicht

Kansrekening en Statistiek

- Geschiedenis
- Loterij
- Toetsen

Email: `broer@math.rug.nl`

URL: `http://www.math.rug.nl/~broer`

Helden I



Jean Pierre de Fermat
(1601-1665)



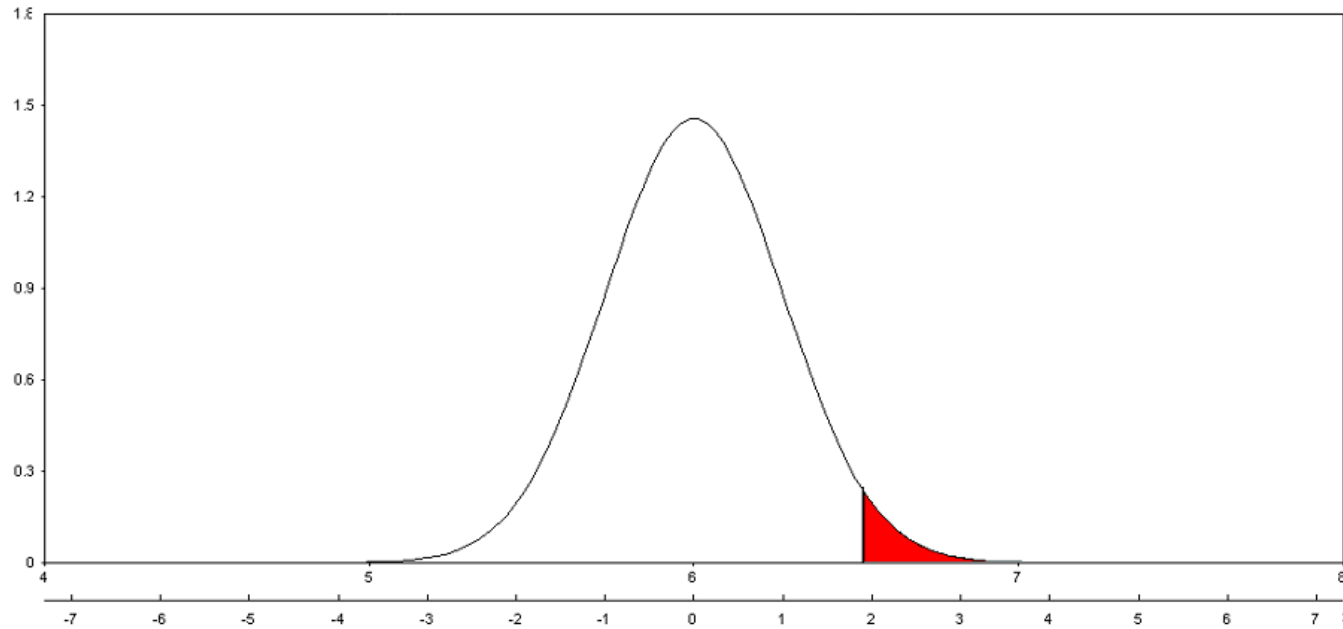
Christiaan Huygens
(1629-1695)

Christiaan HUYGENS, van Rekeningh in Spelen van Geluck, hertaald en toegelicht door
Wim Kleijne, Epsilon Uitgaven 1998

Opzet

- Loterij model:
eindig veel uitkomsten met gelijke kansen.
Tellen combinatoriek.
- Voorwaardelijke kansen
- Toets van een hypothese.
De overschrijdingskans van een uitkomst.

Overschrijdingskans of p -waarde



Overschrijdingskans van een uitkomst
als 'staart' van de kansverdeling

Vindplaatsen van Toeval

- Kansspelen: dobbelen, kaarten, roulette: winstkansen, weddenschapsmodellen
- Steekproef (kijkcijfers)
- Geiger teller
- ...

Systematische beschouwing leidde tot wiskunde

Helden II



Johan de Witt
(1625-1672)



Pierre-Simon Laplace
(1747-1827)

J. de Witt, Waardije van Lijfrenten naar Proportie van Losrenten, 1671

P.-S. Laplace, Théorie Analytique des Probabilités, 1812

Naïef tellen I

- Dobbelsteen, gelijkwaardige uitkomsten

1, 2, 3, 4, 5, 6

elke uitkomst kans $\frac{1}{6}$

- Gebeurtenis $A =$ 'de uitkomst is even',
dan bestaat A uit 2, 4 en 6

Kans $P(A) = \frac{3}{6}$

- Gebeurtenis $A =$ 'geen 6', dan is $P(A) = 1 - \frac{1}{6} = \frac{5}{6}$

Naïef tellen I

- Twee dobbelstenen, gelijkwaardige uitkomsten

(1, 6)	(2, 6)	(3, 6)	(4, 6)	(5, 6)	(6, 6)
(1, 5)	(2, 5)	(3, 5)	(4, 5)	(5, 5)	(6, 5)
(1, 4)	(2, 4)	(3, 4)	(4, 4)	(5, 4)	(6, 4)
(1, 3)	(2, 3)	(3, 3)	(4, 3)	(5, 3)	(6, 3)
(1, 2)	(2, 2)	(3, 2)	(4, 2)	(5, 2)	(6, 2)
(1, 1)	(2, 1)	(3, 1)	(4, 1)	(5, 1)	(6, 1)

Betekenis (i, j) : eerste worp i , tweede j

Elke uitkomst kans $\frac{1}{36}$

- Gebeurtenis $A =$ som der worpen bedraagt 5, dan bestaat uit $(1, 4), (2, 3), (3, 2)$ en $(4, 1)$

Kans $P(A) = \frac{4}{36} = \frac{1}{9}$

Eigenschappen I

- Bestaat A uit de uitkomst $(5, 4)$, dan is
Eerste worp 5:
uitkomsten $(5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6)$ en $(5, 6)$
Tweede worp 4 :
uitkomsten $(1, 4), (2, 4), (3, 4), (4, 4), (5, 4)$ en $(6, 4)$
en $P(A) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$
- Is $A = \text{'geen 6-en'}$ dan is $P(A) = (1 - \frac{1}{6})^2 = \frac{25}{36}$
Wat zit hierachter? Onafhankelijkheid!

Eigenschappen II

- Is $A =$ 'som der worpen groter dan 8' en $B =$ 'dubbel 6', dan is de kans op B gegeven A

$$P(B|A) = \frac{P(A \text{ en } B)}{P(A)}$$

Merk op dat zo

$$P(A \text{ en } B) = P(B|A) \cdot P(A)$$

Tellen levert $P(B|A) = \frac{1}{10}$, *voorwaardelijke kans*

- Algemeen zijn twee gebeurtenissen *onafhankelijk* als $P(A \text{ en } B) = P(A) \cdot P(B)$
Dan geldt dus dat $P(B|A) = P(B)$ en $P(A|B) = P(A)$
- De wiskunde van het toeval is wis en zeker

Schellingen I

Gegeven drie kastjes a , b en c

en elk kastje heeft twee laatjes 1 en 2.

In elk laatje ligt een schelling, een zilveren of een gouden.

Kastje a heeft twee zilveren, kastje b een gouden (lade 2) en een zilveren (lade 1) en kastje c twee gouden schellingen

Iemand kiest aselekt een kastje en daarin aselekt een lade.
Hij vindt een gouden schelling

Wat is de kans dat de andere schelling in ditzelfde kastje ook van goud is?

Schellingen II

Naïeve redenering:

Eerst had je drie kastjes, nu nog slechts twee, het kastje a met de twee zilveren schellingen valt immers af

Van de twee overblijvende b en c is er eentje gunstig, dus is de kans $\frac{1}{2}$

Vraag: Wat is hier fout aan?

Antwoord: Niet de kastjes, maar de laden zijn gelijkwaardig!

Schellingen III

Beter gaat het zo. Gelijkwaardig zijn

$$\begin{array}{ccc} (a, 2) & (b, 2) & (c, 2) \\ (a, 1) & (b, 1) & (c, 1) \end{array}$$

Voor de gebeurtenis

$A =$ 'de gevonden schelling is van goud' geldt

A bestaat uit $(b, 2)$, $(c, 1)$ en $(c, 2)$

Voor de gebeurtenis

$B =$ 'de andere schelling in het kastje is ook van goud'
krijgen we

B bestaat uit $(c, 1)$ en $(c, 2)$

Tellen levert dat $P(B|A) = \frac{2}{3}$

Helden III



Thomas Bayes
(1702-1761)



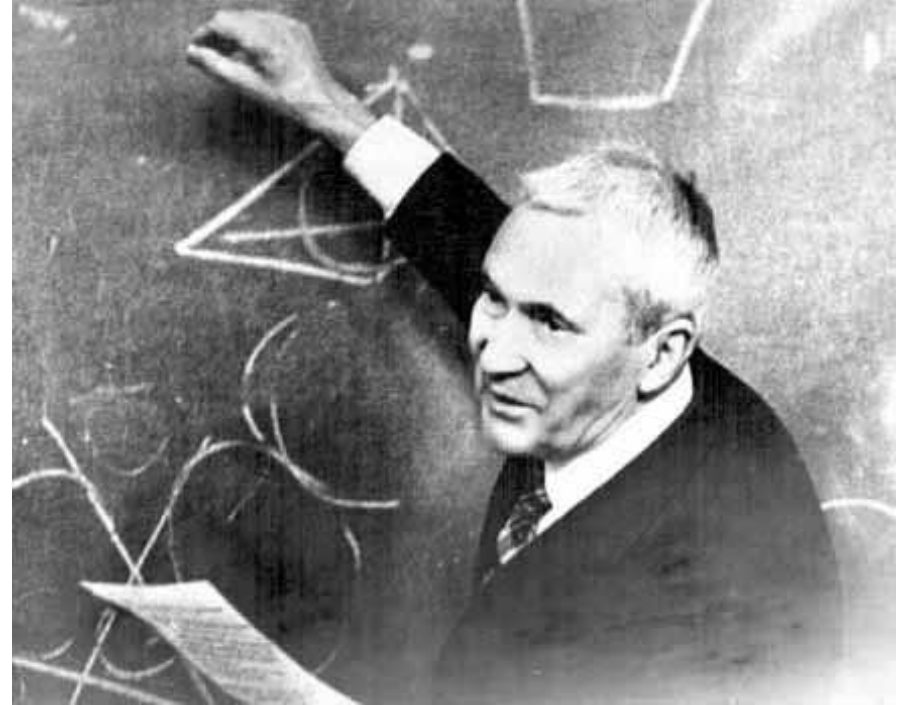
Carl Friedrich Gauß
(1777-1855)

Meer Wiskunde ...

Helden IV



David Hilbert
(1862-1943)



Andrey Nikolaevich Kolmogorov
(1903-1987)

Axiomatische opbouw, moderne theorie

Quizmaster I

Candidaat kiest een der deuren a , b en c kiezen.
Quizmaster opent ANDERE deur waarachter prijs NIET zit.
Hij vraagt “Wil je je keus herzien?”
Wat is de beste strategie voor Candidaat?

Hier zijn gelijkwaardig

(a, c)	(b, c)	(c, c)
(a, b)	(b, b)	(c, b)
(a, a)	(b, a)	(c, a)

Betekenis (x, y) :
de prijs zit achter deur x en
Candidaat kiest eerst y

Quizmaster II

- Is $A = \text{'eerste gok is goed'}$, dan $P(A) = \frac{1}{3}$
In bovenstaande beschrijving bestaat A uit (a, a) , (b, b) en (c, c) en inderdaad $P(A) = \frac{3}{9}$
- Is $B = \text{'eerste gok is verkeerd'}$, dan is $P(B) = 1 - P(A) = \frac{2}{3}$. In bovenstaande beschrijving
 B bestaat uit (b, a) , (c, a) , (c, b) of (a, b) , (a, c) , (b, c)
en inderdaad $P(B) = \frac{6}{9}$
- *Precies* in de gevallen van B is herziening gunstig : QM heeft maar één alternatief en prijs zit achter de nog resterende deur
- Conclusie: herzien is beste strategie, kans op prijs $\frac{2}{3}$

Conclusies I

Opdrachten:

- Generaliseer het QUIZMASTER's probleem naar 4 deuren
- Bedenk loterijmodellen voor het geval Lucia de Berk
- Nu nog een toegift . . .

Twée kinderen I

Iemand krijgt oude vriend op bezoek
Hij heeft inmiddels 2 kinderen,
vriend weet hun geslacht niet

Ze zitten in de woonkamer bijeen
Een dochter steekt hoofd om de deur

Wat is de kans dat het andere kind ook een dochter is?

Twée kinderen II

Spelen met XX en XY chromosomen

↪ gelijkwaardige gezinssamenstellingen

$$\begin{array}{cc} (j, j) & (j, m) \\ (m, j) & (m, m) \end{array}$$

Betekenis (x, y) : eerste kind x , tweede y

Elke uitkomst kans $\frac{1}{4}$

Eerst antwoord: kans is $\frac{1}{3}$

Immers: (j, j) is afgevallen als mogelijkheid

Van de drie overblijvende is precies eentje gunstig

Echter ...

Twée kinderen III

Wat als kinderen hebben geloot wie naar binnen mocht?

Geeft nieuwe gelijkwaardige uitkomsten

$$(j, j, O) \quad (j, m, O)$$

$$(m, j, O) \quad (m, m, O)$$

$$(j, j, J) \quad (j, m, J)$$

$$(m, j, J) \quad (m, m, J)$$

Betekenis (x, y, Z) : gezinssamenstelling is (x, y) en $Z = O$ (udste) of $Z = J$ (ongste) mocht naar binnen.

Nu vallen af: (j, j, O) en (j, j, J) ,
maar ook (j, m, O) en (m, j, J) .

Conclusies II

Van resterende vier

$$(m, j, O), (m, m, O), (j, m, J), (m, m, J)$$

zijn er twee gunstig.

Tweede antwoord: de kans is $\frac{2}{4} = \frac{1}{2}$

- Geslachttonafhankelijke loting met ongelijke kansen geeft ook kans $\frac{1}{2}$
- Antwoorden hangen af van vraagstelling !!
- Kan de quizmaster ook loten?
- Goede proefopzet en modelkeus noodzakelijk

Toetsen I

Iemand werpt 10 maal achtereen met een munt en gooit steeds KOP. Vind je de munt nog wel zuiver?

Loterijmodel zuivere munt:

Gelijkwaardig zijn alle rijtjes met KOP, MUNT ter lengte 10

$(K, M, K, K, M, M, K, M, M, K)$

$(M, M, M, K, K, K, M, M, M, K)$

...

Dat zijn er $2^{10} = 1024$. Dus heeft elke uitkomst A kans

$$P(A) = \frac{1}{1024} \approx 0.001$$

Toetsen II

$$P(10 \text{ maal KOP}) = \frac{1}{1024} \approx 0.001$$

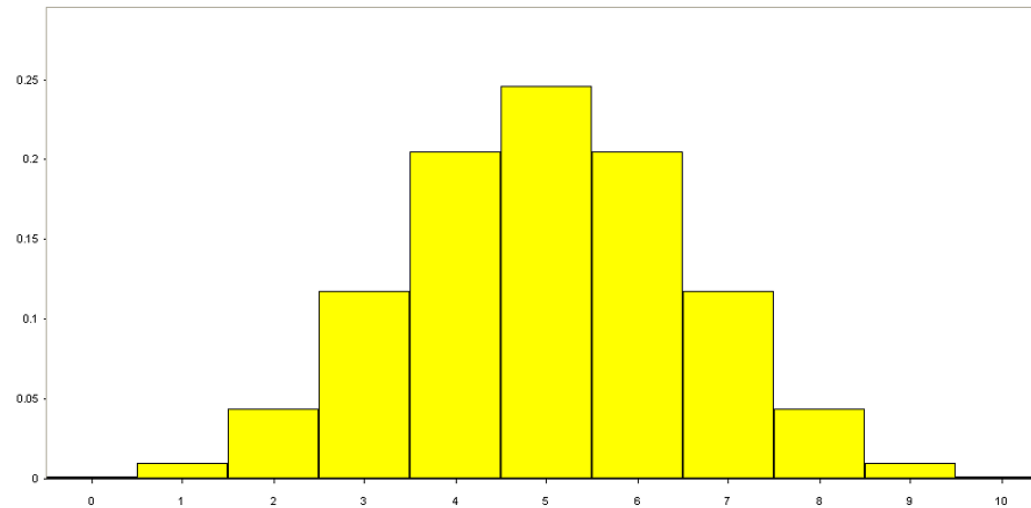
$$P(9 \text{ maal KOP}) = 10 \frac{1}{1024} \approx 0.01 (= 0.0098)$$

$$P(8 \text{ maal KOP}) = \frac{10 \cdot 9}{1 \cdot 2} \frac{1}{1024} \approx 0.045$$

Dus: Overschrijdingskans '9 maal KOP' =

$$= 0.0098 + 0.001 = 0.0108 \approx 1.1\%$$

Toetsen III



Kansverdeling van het aantal keren KOP bij 10 worpen

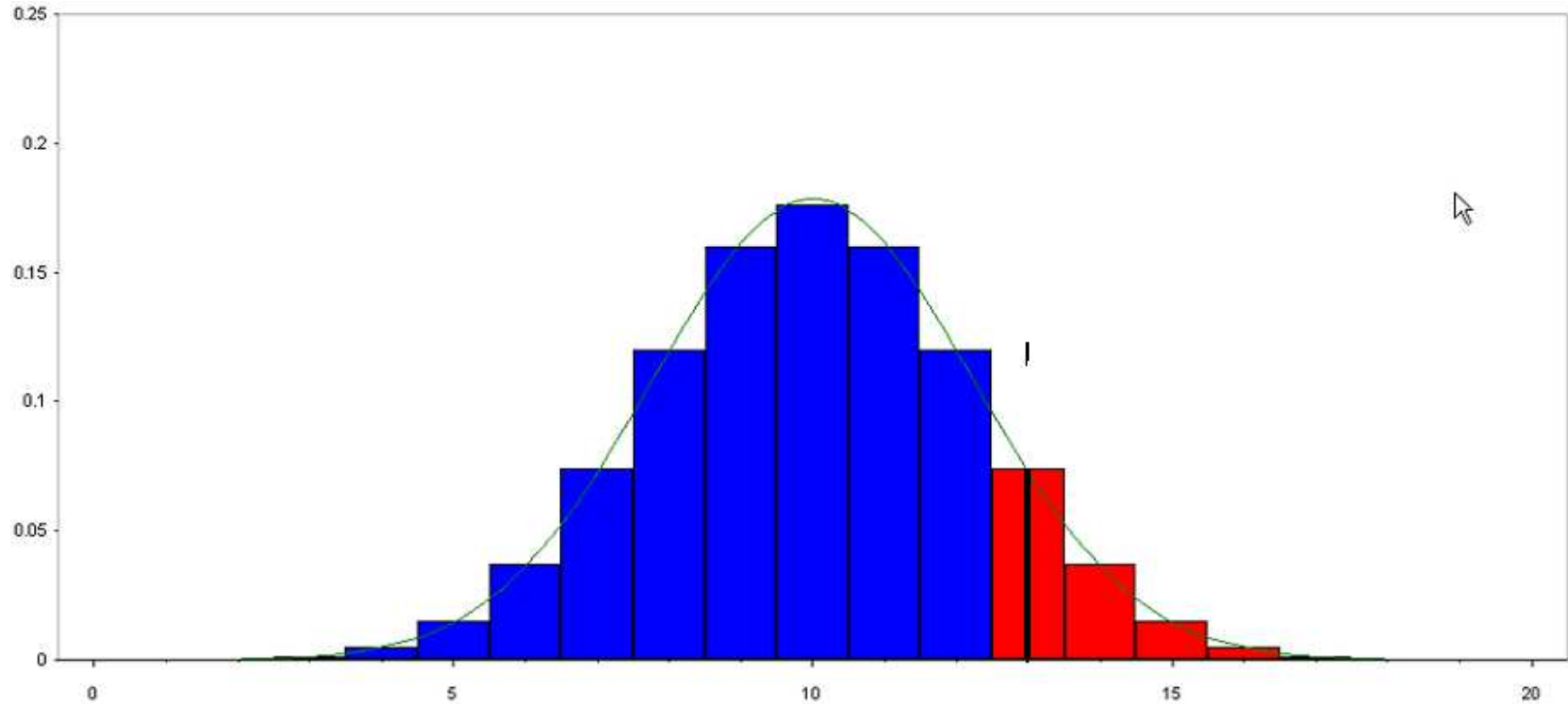
Overschrijdingskansen bij uitkomsten

$$10 \text{ maal KOP : } p = 0.1\%$$

$$9 \text{ maal KOP : } p = 0.1 + 1.0 = 1.1\%$$

$$8 \text{ maal KOP : } p = 0.1 + 1.0 + 4.5 = 5.6\%$$

Toetsen IV



Kansverdeling van het aantal keren KOP bij 20 worpen

Conclusies

Wanneer accepteer je de *hypothese* dat de munt zuiver is, nog als redelijk d.w.z., als niet *te* onwaarschijnlijk?

Dit is een kwestie van afspraak.

Men hanteert een overschrijdings-*drempel*:

5% of 1% of 0.1%

Is de overschrijdingskans lager dan deze drempel dan *verwerpt* men de hypothese.

Opdrachten:

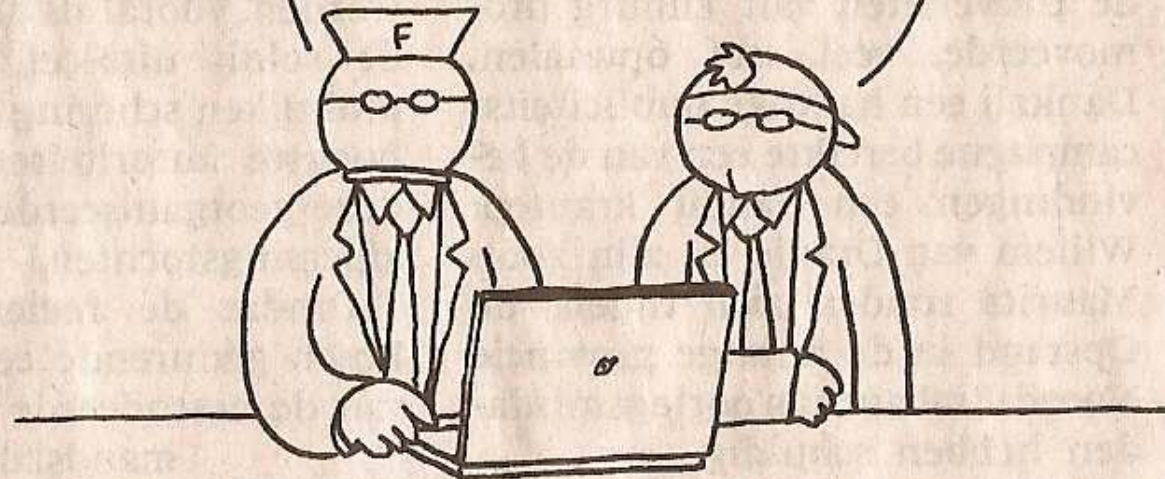
- In welke van bovenstaande gevallen moet de hypothese 'de munt is zuiver' verworpen?
- Pas bovenstaande aan voor een dobbelsteen die veel 6-sen te zien geeft.
- Formuleer bij Lucia de Berk een relevant toetsingsprobleem.

Fokke en Sukke

FOKKE & SUKKE
VOELEN DAT AAN HUN WATER

DE KANS DAT VRIJWEL ALLE
HOOGLERAREN STATISTIEK HET
EENS ZIJN

...IS NATUURLIJK
HEEL ERG KLEIN.



NRC-HL 2-11-07

RGvT

www.foksuk.nl

Literatuur

1. K.L. Chung and F. AitSahlia, *Elementary Probability Theory*, Springer Verlag 2003
2. Larry Gonick en Woollcot Smith, *Het Stripverhaal van de Statistiek*, Epsilon Uitgaven 1994
3. *Euclides* 4, Februari 2008