ZW

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM
AFDELING ZUIVERE WISKUNDE

Syllabus van het oriënterend colloquium

Waarschijnlijkheidsrekening

cursus 1968-1969

ZW

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

Colloquium "Waarschijnlijkheidsrekening" (1968-1969)

4 september 1968

Spreker: J. van de Lune

1. Inleiding

In dit colloquium zullen we een aantal hoofdstukken behandelen uit dat gedeelte van de zuivere wiskunde dat waarschijnlijkheidsrekening (WHR) heet.

De WHR bestond enkele honderden jaren geleden voornamelijk uit (min of meer exacte) beschouwingen over de mogelijkheden tot het maken van winst bij diverse zogenaamde kansspelen. Een in dit verband haast legendarische figuur is wel Chevalier de Meré. Op grond van bepaalde overwegingen meende hij dat het voordelig was om bij 24 worpen met twee dobbelstenen te wedden op minstens één maal dubbelzes.

In de praktijk bleek evenwel dat hij met deze strategie vaker verloor dan won. Naar aanleiding hiervan ging hij zich bij zijn vriend Blaise Pascal beklagen over de wiskunde. Als antwoord op dit beklag rekende Pascal hem voor dat het, in overeenstemming met de ervaring, onvoordelig is te wedden op minstens een maal dubbelzes in 24 worpen met twee dobbelstenen.

Men zegt wel dat Pascal bij deze gelegenheid de WHR uitvond.

De manier waarop Pascal de kansspelen behandelde mag stellig exact genoemd worden, ook al ging hij niet axiomatisch te werk. Een moderne wiskundige behandeling moet namelijk axiomatisch zijn.

Onze eerste doelstelling is dan ook een zuiver wiskundig systeem te creëren dat voor zeel veel toevalsprocessen als mathematisch model kan dienen. In de volgende paragraaf zullen wij aan de hand van een eenvoudig voorbeeld laten zien hoe dit in zijn werk gaat.

2. Als toevalsproces nemen wij het éénmaal werpen met één dobbelsteen. Het is duidelijk dat het resultaat van dit toevalsproces (experiment) niet met zekerheid te voorspellen is. Na enige idealisering kan men wel met zekerheid zeggen welke resultaten mogelijk zijn. In het algemeen zal elk toevalsproces een welomschreven collectie van mogelijke resultaten (uitkomsten) met zich meebrengen, zodanig dat twee verschillende resultaten nooit gelijktijdig kunnen voorkomen. We kunnen dus zeggen dat elk toevalsproces aanleiding geeft tot een verzameling van mogelijke enkelvoudige uitkomsten. In ons voorbeeld hebben we te maken met de verzameling van alle mogelijke aantallen ogen die geworpen kunnen worden, dus met de verzameling $\{1, 2, 3, 4, 5, 6\}$.

   Ons streven naar axiomatisering van de WHR haakt op deze beschouwing in door als grondsysteem voor de WHR een abstracte niet lege verzameling E te kiezen, die, als men wil, geïnterpreteerd dient te worden als de verzameling van alle mogelijke enkelvoudige uitkomsten van een of ander experiment.

   In dit colloquium zal de verzameling E steeds hoogstens aftelbaar zijn. In een uiterst algemene opzet van de WHR stelt men aan E nauwelijks enige voorwaarde, maar dit heeft dan wel tot gevolg dat de verdere mathematische behandeling heel wat gecompliceerder wordt dan in dit colloquium gewenst zou zijn. We keren terug naar ons dobbelsteen voorbeeld waar $E = \{1, 2, 3, 4, 5, 6\}$.

   We zijn gewoon aan elk van de mogelijke enkelvoudige uitkomsten een zekere kans toe te kennen; in het geval van een zogenaamde "zuivere"dobbelsteen zijn de kansen op alle mogelijke uitkomsten gelijk aan $\frac{1}{6}$. Maar ook in het geval van een "onzuivere" dobbelsteen zal men aan elk der mogelijke uitkomsten een kans willen toekennen. Over de numerieke waarden van deze kansen kan men in het onzekere verkeren, maar men zal het er over eens zijn dat elke kans een niet negatief reëel getal tussen 0 en 1 is.

   Verder zal men het er over eens zijn dat de kansen op alle mogelijke enkelvoudige uitkomsten (behorende bij een en hetzelfde experiment) tot 1 sommeren.

Aansluitend hierop nemen we, in ons streven naar axiomatisering van de WHR, aan dat op E een funktie p gedefinieerd is met de volgende eigenschappen:

Duiden we E aan met $\{E_1, E_2, E_3, \ldots\}$
dan is

1) $p : E \to \mathbb{R}$ (= verzameling van alle reële getallen)

2) $p(E_i) \geq 0$, $(i = 1, 2, 3, \ldots)$

3) $\sum_{i=1}^{\infty} p(E_i) = 1$.

We zeggen dat het paar (E, p) een <u>waarschijnlijkheidsveld</u> is als p aan bovenstaande voorwaarden voldoet.

<u>Voorbeeld 1.</u> Het éénmaal werpen met een zuivere dobbelsteen (intuïtief geïnterpreteerd) geeft aanleiding tot het volgende (zuiver mathematisch geformuleerde) waarschijnlijkheidsveld:

$E = \{1, 2, 3, 4, 5, 6\}$ met

$p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = \frac{1}{6}$.

<u>Voorbeeld 2.</u> Het éénmaal werpen met een zuivere munt geeft aanleiding tot het volgende WH-veld:

$E = \{K, M\}$, (waarbij K staat voor "kruis" en M voor "munt")

met $p(K) = p(M) = \frac{1}{2}$.

<u>Voorbeeld 3.</u> Het éénmaal werpen met een punaise kan als resultaat hebben

P (= punt,) of R (= rug, $\perp$).

Het bijbehorende mathematische model is

$$E = \{P, R\} \text{ met } p(P) = k_1 \text{ en } p(R) = k_2$$

waarbij $\quad k_1, k_2 \geqq 0 \quad \text{en} \quad k_1 + k_2 = 1.$

Het zal blijken dat de veronderstelling $k_1 = k_2 = \frac{1}{2}$ doorgaans moeilijk te handhaven is. In het extreme geval dat de punaise de gedaante van een lange spijker heeft ligt het zeer voor de hand te veronderstellen dat $k_1 > k_2$.

Voorbeeld 4. We werpen tweemaal met een zuivere munt. Het bijbehorende WH-veld is dan

$$E = \{(K, K), (K, M), (M, K), (M, M)\} \text{ met}$$

$$p(K, K) = p(K, M) = p(M, K) = p(M, M) = \frac{1}{4}.$$

N.B. Kortheidshalve schrijven we $p(K, K)$ in plaats van $p((K, K))$.

Voorbeeld 5. We werpen weer tweemaal met een zuivere munt en zijn daarbij geïnteresseerd in het aantal malen "kruis" dat geworpen wordt. Als mathematisch model voor dit experiment zou men kunnen nemen:

$$E = \{0, 1, 2\} \text{ met } p(0) = p(2) = \frac{1}{4} \text{ en } p(1) = \frac{1}{2}.$$

3. Gebeurtenissen. (eventualiteiten)

Zij (E, p) een WH-veld met $E = \{E_1, E_2, E_3, \ldots\}$.

Definitie. Een deelverzameling A van E heet een gebeurtenis. De kans op de gebeurtenis A (notatie: $P\{A\}$) wordt gedefinieerd als

$$P\{A\} = \sum_{E_i \in A} p(E_i) \text{ als } A \neq \emptyset$$

en $\quad P\{\emptyset\} = 0.$

E heet de zekere gebeurtenis en $\emptyset$ de onmogelijke gebeurtenis.
De punten E_i van E heten enkelvoudige gebeurtenissen.
Met behulp van deze definitie en een aantal eenvoudige eigenschappen van convergente reeksen met positieve termen bewijst men de volgende drie stellingen.

Stelling 1. Voor elke gebeurtenis A geldt

$$0 \leq P\{A\} \leq 1.$$

Stelling 2. Indien we het complement van de gebeurtenis A aangeven met $\neg A$ (de negatie van A) dan geldt

$$P\{\neg A\} = 1 - P\{A\}.$$

Anders geschreven $\quad P\{A\} + P\{\neg A\} = 1.$

Opmerking. In de literatuur schrijft men voor $\neg A$ vaak $\bar{A}$.

Stelling 3. Zijn A en B gebeurtenissen dan is

$$P\{A \cup B\} = P\{A\} + P\{B\} - P\{A \cap B\}.$$

We kunnen dit ook schrijven als

$$P\{A \cup B\} + P\{A \cap B\} = P\{A\} + P\{B\}.$$

Opmerking. De gebeurtenis $A \cup B$ heet de disjunctie, de gebeurtenis $A \cap B$ de conjunctie van A en B.

Gevolg. Als $A \cap B = \emptyset$ dan geldt

$$P\{A \cup B\} = P\{A\} + P\{B\}.$$

Met gaat ook gemakkelijk na dat, als $A = \bigcup_{i=1}^{\infty} A_i$ met $A_i \cap A_j = \emptyset$ voor $i \neq j$ (A_i en A_j zijn voor $i \neq j$ elkaar uitsluitende gebeurtenissen),

$$P\{A\} = \sum_{i=1}^{\infty} P\{A_i\}.$$

Opgave. Als $A = \bigcup_{i=1}^{\infty} A_i$ dan geldt

$$P\{A\} \leq \sum_{i=1}^{\infty} P\{A_i\} \text{ (ongelijkheid van Boole).}$$

Hierbij behoeven A_i en A_j voor $i \neq j$ elkaar niet uit te sluiten.

Stelling 4. Als $A = \bigcup_{i=1}^{n} A_i$ dan is

$$P\{A\} = \sum_{i=1}^{n} P\{A_i\} - \sum_{i<j\leq n} P\{A_i \cap A_j\} + \sum_{i<j<k\leq n} P\{A_i \cap A_j \cap A_k\} +$$

$$+ \dots + (-1)^{n-1} P\{A_1 \cap A_2 \cap \dots \cap A_n\}.$$

Bewijs: De stelling is juist voor $n = 1$ (en volgens stelling 2 ook voor $n = 2$). Neem aan dat de stelling nog juist is voor $n = m$. We tonen aan dat de stelling dan ook nog juist is voor $n = m + 1$.

$$P\{A_1 \cup A_2 \cup \dots \cup A_m \cup A_{m+1}\} = P\{(A_1 \cup A_2 \cup \dots \cup A_m) \cup A_{m+1}\} =$$

$$= P\{A_1 \cup A_2 \cup \dots \cup A_m\} + P\{A_{m+1}\} - P\{(A_1 \cup A_2 \cup \dots \cup A_m) \cap A_{m+1}\} =$$

$$= P\{A_1 \cup A_2 \cup \dots \cup A_m\} + P\{A_{m+1}\} - P\{(A_1 \cap A_{m+1}) \cup (A_2 \cap A_{m+1}) \cup \dots \cup (A_m \cap A_{m+1})\} =$$

$$= \sum_{i=1}^{m} P\{A_i\} - \sum_{i<j\leq m} P\{A_i \cap A_j\} + \sum_{i<j<k\leq m} P\{A_i \cap A_j \cap A_k\} - + \dots$$

$$+ P\{A_{m+1}\} - \sum_{i=1}^{m} P\{A_i \cap A_{m+1}\} + \sum_{i<j\leq m} P\{A_i \cap A_j \cap A_{m+1}\} - + \dots =$$

$$= \sum_{i=1}^{m+1} P\{A_i\} - \sum_{i<j\leq m+1} P\{A_i \cap A_j\} + \sum_{i<j<k\leq m+1} P\{A_i \cap A_j \cap A_k\} - + \dots$$

Hiermee is de stelling door volledige inductie bewezen.

De juistheid van stelling 4 is ook als volgt gemakkelijk in te zien. In de opgegeven formule voor $P\{A\}$ spelen alleen kansen $p(E_i)$ met $E_i \in A$ een rol. Van elk van deze kansen bepalen we de coëfficient. Zij E_i een element van A dat in precies r deelverzamelingen A_j ligt; dan is op grond van eenvoudige combinatorische overwegingen de coëfficient van $p(E_i)$ gelijk aan

$$\binom{r}{1} - \binom{r}{2} + \binom{r}{3} - + \ldots + (-1)^{r-1}\binom{r}{r}.$$

Het binomium van Newton

$$(a+b)^r = \sum_{k=0}^{r} \binom{r}{k} a^k b^{r-k}$$

levert voor a = -1 en b = +1,

$$0 = 0^r = (-1 + 1)^r = \sum_{k=0}^{r} \binom{r}{k} (-1)^k =$$

$$= 1 - \binom{r}{1} + \binom{r}{2} - + \ldots + (-1)^r \binom{r}{r},$$

met als gevolg

$$\binom{r}{1} - \binom{r}{2} + \ldots + (-1)^{r-1}\binom{r}{r} = 1.$$

De gezochte coëfficient is dus 1 en is dus onafhankelijk van r. Hieruit volgt dat elke $p(E_i)$ met $E_i \in A$ de coëfficient 1 heeft. De opgegeven formule voor $P\{A\}$ is dus gelijkwaardig met

$$\sum_{E_i \in A} p(E_i) = P\{A\}.$$

Opgaven

1. Men werpt tweemaal met een zuivere munt. Hoe groot is de kans dat

    a de eerste worp "kruis" is,

    b de tweede worp "kruis" is,

    c de eerste èn de tweede worp "kruis" is,

    d de eerste of de tweede worp "kruis" is,

    e er precies één maal munt wordt geworpen?

2. Uit de getallen 1, 2, 3, 4 en 5 wordt aselect een getal gekozen en terzijde gelegd; uit de resterende getallen trekt men daarna aselect nog een getal. Hoe groot is de kans dat

    a de eerste trekking even is,

    b de eerste trekking oneven is,

    c de tweede trekking even is,

    d de tweede trekking oneven is,

    e beide trekkingen even zijn,

    f beide trekkingen oneven zijn,

    g de eerste trekking ≤ 4 èn de tweede trekking ≥ 3 is,

    h de twee trekkingen een g.g.d. > 1 hebben,

    i de som van de trekkingen ≥ 7 is,

    j het product van de trekkingen priem is?

3. Men werpt tweemaal met een zuivere dobbelsteen. Hoe groot is de kans dat de som van de aantallen geworpen ogen gelijk is aan n ($n = 0, 1, 2, 3, \ldots, 13$)?
    Zet in een grafiek deze kansen uit tegen n.

4. Men werpt driemaal met een zuivere dobbelsteen. Hoe groot is de kans dat de som van het aantal geworpen ogen gelijk is aan n ($n = 0, 1, 2, 3, \ldots, 20$)? Zet in een grafie deze kansen uit tegen n.

5. Men werpt n maal met een zuivere munt. Hoe groot is de kans dat er precies k ($0 \leq k \leq n$) maal kruis wordt geworpen?

6. Men wijst aselect n mensen aan. Hoe groot is de kans dat minstens twee van deze mensen op dezelfde dag jarig zijn?

7. N heren hebben tijdens een vergadering hun hoed aan een kapstok gehangen. Als deze heren na afloop van de vergadering aselect een hoed van de kapstok nemen, hoe groot is dan de kans dat minstens een van hen zijn eigen hoed heeft genomen?

8. Men schrijft aselect k cijfers op. Hoe groot is de kans dat het cijfer 7 vijfmaal voorkomt?

9. Men werpt viermaal met een zuivere dobbelsteen. Hoe groot is de kans op minstens eenmaal 6?

10. Men werpt 24 maal met twee dobbelstenen. Hoe groot is de kans op minstens eenmaal dubbelzes?

11. Op een roulette in Monte Carlo staan de getallen 0 t.e.m. 36. Men zegt dat er eens een serie van 28 maal "even" zou zijn voorgekomen. Hoe groot is de kans op deze gebeurtenis?

4. Voorwaardelijke kansen.

We gaan uit van een WH-veld (E, p) met $E = \{E_1, E_2, E_3, \ldots\}$.
Zij $H \subset E$ zodanig dat $P\{H\} > 0$.
Voor een punt $E_i \in H$ definiëren we

$$p^{*}(E_i) = \frac{p(E_i)}{P\{H\}} \ .$$

Dan is (H, p^{*}) een WH-veld.

<u>Definitie:</u> Zij $A \subset E$ en $P\{H\} > 0$. Dan definieren we "<u>de kans op A onder de voorwaarde H</u>" (notatie $P\{A \mid H\}$) als

$$P\{A \mid H\} = \sum_{E_i \in A \cap H} p^{*}(E_i).$$

<u>Stelling 1.</u> $P\{A \mid H\} = \frac{P\{A \cap H\}}{P\{H\}}$

Bewijs: $$P\{A \mid H\} = \sum_{E_i \in A \cap H} p^{*}(E_i) = \sum_{E_i \in A \cap H} \frac{p(E_i)}{P\{H\}} =$$

$$= \frac{1}{P\{H\}} \cdot \sum_{E_i \in A \cap H} p(E_i) = \frac{P\{A \cap H\}}{P\{H\}} \ .$$

Dit resultaat kan ook geschreven worden als $P\{A \cap H\} = P\{A \mid H\}.P\{H\}$.

<u>Stelling 2.</u> Zij A een gebeurtenis in E en $H_1, H_2, \ldots, H_n, \ldots$ een eindig of aftelbaar stelsel gebeurtenissen in E zodanig dat 1) $H_i \cap H_j = \emptyset$ als $i \neq j$, 2) $E = \bigcup_{i=1}^{\infty} H_i$, 3) $P\{H_i\} > 0$ voor alle i.

Dan geldt:

$$P\{A\} = \sum_{i=1}^{\infty} P\{A \mid H_i\}. P\{H_i\}.$$

Bewijs:

$$P\{A\} = P\{A \cap E\} = P\{A \cap (\bigcup_{i=1}^{\infty} H_i)\} =$$

$$= P\{\bigcup_{i=1}^{\infty} (A \cap H_i)\} = \sum_{i=1}^{\infty} P\{A \cap H_i\} =$$

$$= \sum_{i=1}^{\infty} P\{A \mid H_i\}. P\{H_i\}.$$

Opmerking. De voorwaarde $E = \bigcup_{i=1}^{\infty} H_i$ kan vervangen worden door $A \subset \bigcup_{i=1}^{\infty} H_i$.

Opgaven.

1. We hebben de beschikking over een zuivere munt en een zuivere dobbelsteen. Op de munt is op de "kruis" kant een 0 en op de "munt" kant een 1 aangebracht. We maken eerst een aselecte keuze uit de collectie "kruis, dobbelsteen" en werpen daarna met het gekozen object één maal. Hoe groot is de kans dat er een 1 geworpen wordt?

2. We hebben de beschikking over N vazen $V_1, V_2, \ldots, V_n$. In de vaas V_i zitten i witte en N-i zwarte knikkers (i = 1, 2, ... , N). Eerst wordt aselect een vaas aangewezen; daarna worden uit de gekozen vaas aselect met teruglegging twee knikkers gekozen. Hoe groot is de kans dat beide knikkers wit zijn?

3. We hebben weer de beschikking over N vazen zoals in vraagstuk 2 is aangegeven. We wijzen eerst aselect een vaas aan en nemen uit de gekozen vaas aselect en met teruglegging n knikkers. Als al deze knikkers wit zijn hoe groot is dan de kans dat er bij nog een trekking uit dezelfde vaas weer een witte knikker te voorschijn komt?

5. Als $P\{H\} > 0$, toon dan aan dat

a $P\{\emptyset \mid H\} = 0, \quad P\{E \mid H\} = 1, \quad 0 \leq P\{A \mid H\} \leq 1$

b $P\{(\neg A) \mid H\} = 1 - P\{A \mid H\}$

c $P\{(A \cup B) \mid H\} = P\{A \mid H\} + P\{B \mid H\} - P\{(A \cap B) \mid H\}$

d Als $A = \bigcup_{i=1}^{\infty} A_i$ dan is

$$P\{A \mid H\} \leq \sum_{i=1}^{\infty} P\{A_i \mid H\}$$

e Als $A = \bigcup_{i=1}^{n} A_i$ dan is

$$P\{A \mid H\} = \sum_{i=1}^{n} P\{A_i \mid H\} - \sum_{i<j\leq n} P\{(A_i \cap A_j) \mid H\} +$$

$$+ \sum_{i<j<k\leq n} P\{(A_i \cap A_j \cap A_k) \mid H\} - + \ldots$$

f Als $A \subset B$ en $P\{A\} > 0$ dan is

$$P\{B \mid A\} = 1$$

g Als $A \subset B$ en $P\{B\} > 0$ dan is

$$P\{A \mid B\} = \frac{P\{A\}}{P\{B\}} .$$

h Als A en B elkaar uitsluitende gebeurtenissen zijn en $P\{B\} > 0$ dan is

$$P\{A \mid B\} = 0$$

5. Onafhankelijkheid van gebeurtenissen.

Zij $H \subset E$ met $P\{H\} > 0$. In vele gevallen kan men constateren dat voor zekere gebeurtenissen A geldt

$$P\{A \mid H\} \neq P\{A\}.$$

Intuïtief gesproken wil dit zeggen dat de informatie betreffende het optreden van de gebeurtenis H het optreden van een gebeurtenis A kan beïnvloeden. Vergelijken we het percentage van de gehele Nederlandse bevolking dat aan zekere ziekte A lijdt met het percentage van de tegen ziekte A ingeënte Nederlanders dat aan ziekte A lijdt dan zal heel vaak het eerste percentage groter zijn dan het tweede.
We zouden kunnen definiëren dat de gebeurtenis A afhankelijk is van de gebeurtenis H (met $P\{H\} > 0$) als $P\{A \mid H\} \neq P\{A\}$ en dat A onafhankelijk is van H als

$$P\{A \mid H\} = P\{A\}.$$

Voor $P\{A \mid H\} \neq P\{A\}$ kunnen we schrijven

$$P\{A \cap H\} \neq P\{A\}.\ P\{H\}.$$

en $P\{A \mid H\} = P\{A\}$ is gelijkwaardig met

$$P\{A \cap H\} = P\{A\}.\ P\{H\}.$$

Naar aanleiding hiervan geven we de volgende

Definitie: De gebeurtenis A heet stochastisch onafhankelijk van de gebeurtenis H als

$$P\{A \cap H\} = P\{A\}.\ P\{H\}.$$

en stochastisch afhankelijk van H als $P\{A \cap H\} \neq P\{A\}\ P\{H\}$.

Opmerking. In deze definitie wordt niet geëist dat $P\{H\} \neq 0$.

Gevolg: Als A (on)afhankelijk is van H dan is ook H (on)afhankelijk van A. We kunnen dus zonder meer spreken over twee (on)afhankelijke gebeurtenissen.

Stelling 1. Als A en H onafhankelijk zijn dan zijn ook de volgende paren gebeurtenissen onafhankelijk

$$(\neg A, H)\ ,\ (A, \neg H),\ (\neg A, \neg H).$$

Bewijs:

$$P\{H\} = P\{E \cap H\} = P\{(\neg A) \cup A) \cap H\} =$$

$$= P\{((\neg A) \cap H) \cup (A \cap H)\} = P\{(\neg A) \cap H\} + P\{A \cap H\}$$

$$= P\{(\neg A) \cap H\} + P\{A\}.\ P\{H\}$$

Dus: $P\{(\neg A) \cap H\} = P\{H\}\ (1 - P\{A\}) = P\{\neg A\}.\ P\{H\}.$

De rest van het bewijs wordt als opgave aan de lezer overgelaten.

Opgaven.

1. We werpen tweemaal met een zuivere munt. A zij de gebeurtenis "de eerste worp is kruis" en B de gebeurtenis "de tweede worp is munt".
   Zijn A en B onafhankelijk?

2. Uit de getallen 1, 2, 3 en 4 worden aselect zonder teruglegging twee getallen gekozen. A zij de gebeurtenis "de eerste trekking is even", B de gebeurtenis "de tweede trekking is even". Zijn A en B onafhankelijk?

3. We werpen tweemaal met een zuivere dobbelsteen. A is de gebeurtenis dat het aantal ogen van de eerste worp ≤ 5 is. B is de gebeurtenis dat de som der aantallen ogen van beide worpen gelijk is aan 8. Zijn A en B onafhankelijk?

6. Stochastische funkties (stochastische variabelen)

Definitie. Een stochastische funktie is een reëelwaardige funktie op een WH-veld (E, p).
Stochastische funkties zullen we doorgaans aanduiden met de letters U, V, W, X, Y en Z, eventueel voorzien van de nodige indices.

Definitie. Indien X een stochastische funktie is op (E, p) en de reeks

$$\sum_{i=1}^{\infty} X(E_i)\, p(E_i)$$

is absoluut convergent, dan heet de som van deze reeks de (mathematische) verwachting van X.
Notatie:

$$\mathcal{E}(X) = \sum_{i=1}^{\infty} X(E_i)\, p(E_i) .$$

Voor $\mathcal{E}(X)$ schrijven we ook vaak μ_X.

Opmerking. Men spreekt alleen van $\mathcal{E}(X)$ als de reeks $\sum_{i=1}^{\infty} X(E_i).p(E_i)$ absoluut convergeert. Dit wordt voornamelijk gedaan om $\mathcal{E}(X)$ niet afhankelijk te doen zijn van de nummering van het WH-veld $E = \{E_1, E_2, E_3, \ldots\}$.

Stelling 1. Een stochastische funktie die op geheel E de constante waarde a aanneemt heeft de verwachting a.

Bewijs: Dit volgt onmiddellijk uit de definities.

Stelling 2. Indien X een stochastische funktie is met verwachting $\mathcal{E}(X)$ dan is voor elke constante a

$$\mathcal{E}(a\,X) = a.\mathcal{E}(X).$$

Bewijs: $\sum_{i=1}^{\infty} (a\,X)\,(E_i).\ p(E_i) = \sum_{i=1}^{\infty} a.\ X(E_i).\ p(E_i) = a.\mathcal{E}(X).$

Ga na dat $\mathcal{E}(a\,X)$ inderdaad bestaat.

Stelling 3. Zijn A en Y stochastische funkties met verwachtingen $\mathcal{E}(X)$ en $\mathcal{E}(Y)$ dan bestaat ook de verwachting van X + Y en er geldt

$$\mathcal{E}(X + Y) = \mathcal{E}(X) + \mathcal{E}(Y).$$

Bewijs: Dat $\mathcal{E}(X + Y)$ bestaat berust op de ongelijkheid

$$|x + y| \leqq |x| + |y|.$$

Verder is $\mathcal{E}(X + Y) = \sum_{i=1}^{\infty} (X + Y)\,(E_i).\ p(E_i) =$

$$= \sum_{i=1}^{\infty} \{X(E_i) + Y(E_i)\}.\ p(E_i) = \sum_{i=1}^{\infty} X(E_i).\ p(E_i) + \sum_{i=1}^{\infty} Y(E_i).\ p(E_i) =$$

$$= \mathcal{E}(X) + \mathcal{E}(Y).$$

Opmerking. $\mathcal{E}$ kan dus worden opgevat als een lineaire funktionaal op de lineaire ruimte van alle stochastische funkties X waarvoor $\sum_{i=1}^{\infty} X(E_i).\ p(E_i)$ absoluut convergeert.

<u>Stelling 4.</u> Zij X een stochastische funktie waarvan $\mathcal{E}(X^n)$, $(n = 1, 2, 3, \ldots)$ bestaat. Dan bestaat ook $\mathcal{E}(X^{n-1})$.

Bewijs: Dit berust op de volgende ongelijkheid

$$x^{n-1} < x^n + 1 \text{ als } x \geqq 0.$$

<u>Definitie.</u> Indien X een stochastische funktie is zodanig dat $\mathcal{E}(X^2)$ bestaat, dan bestaat ook

$$\mathcal{E}(X - \mathcal{E}(X))^2 \qquad \text{(ga dit na),}$$

en deze grootheid heet de <u>variantie</u> van X.
Notatie: Var(X) = $(X - \mu_X)^2$.
De vierkantswortel uit Var(X) heet de <u>spreiding</u> of ook wel de <u>standaard-afwijking</u> van X.
Voor de <u>spreiding</u> van X schrijven we vaak σ_X (of kortweg σ als misverstand is uitgesloten) zodat Var(X) = σ_X^2.

<u>Stelling 5.</u> $\mathrm{Var}(X) = \mathcal{E}(X^2) - \mu_X^2$.

Bewijs: $\mathrm{Var}(X) = \mathcal{E}(X - \mu_X)^2 = \mathcal{E}(X^2 - 2\,\mu_X . X + \mu_X^2) =$

$$= \mathcal{E}(X^2) - 2\mu_X . \mathcal{E}(X) + \mu_X^2 = (X^2) - \mu_X^2.$$

<u>Stelling 6.</u>
$$\left.\begin{array}{l} \mathrm{Var}\ (a\,X) = a^2 .\ \mathrm{Var}\ (X), \\ \\ \mathrm{Var}\ (X + a) = \mathrm{Var}\ (X) \end{array}\right\} \text{(a constant).}$$

Bewijs: Volgt onmiddellijk uit de definities.

Opgaven.

1. Zij $E = \{0, 1\}$, $p(0) = \alpha$ en $p(1) = 1 - \alpha = \beta$.
   Bereken van de volgende stochastische funkties de verwachting en de variantie

   a $X: E \to R$ met $X(0) = 1$ en $X(1) = 0$.

   b $Y: E \to R$ met $Y(0) = 0$ en $Y(1) = 1$.

2. Zij $E = \{0, 1, 2, 3, \ldots\}$ met $p(i) = e^{-\lambda} \cdot \frac{\lambda^i}{i!}$, $(i = 0, 1, 2, 3, \ldots)$.
   Als $X(i) = i$ voor alle $i \in E$ bereken dan

$$\mathcal{E}(X) \text{ en } \operatorname{Var}(X).$$

3. Zij $E = \{0, 1, 2, 3, \ldots , n\}$ met $p(i) = \binom{n}{i} \alpha^i (1-\alpha)^{n-1}$, $(0<\alpha<1)$.

   Als $X_1(i) = i$ en $X_2(i) = i^2$ bereken dan de verwachting en de variantie van X_1 en X_2.

4. Wat kan men concluderen uit $\operatorname{Var}(X) = 0$?

Zij X een stochastische funktie met verwachting μ en spreiding $\sigma \neq 0$. Dan heet de funktie

$$X^* = \frac{X - \mu}{\sigma}$$

de bij X behorende genormaliseerde stochastische funktie.

Stelling 7. $\mathcal{E}(X^*) = 0$ en $\operatorname{Var}(X^*) = 1$.

Bewijs:
$$\mathcal{E}(X^*) = \mathcal{E}\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}\, \mathcal{E}(X - \mu) =$$

$$= \frac{1}{\sigma} \cdot \{\mathcal{E}(X) - \mu\} = 0,$$

$$\operatorname{Var}(X^*) = \operatorname{Var}\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2} \cdot \operatorname{Var}(X - \mu) =$$

$$= \frac{1}{\sigma^2} \cdot \mathrm{Var}\ (X) = 1.$$

<u>Definitie.</u> Zijn X en Y stochastische funkties met verwachtingen μ_X en μ_Y dan wordt de <u>covariantie van X en Y</u> (notatie Cov (X, Y)) gedefinieerd door

$$\mathrm{Cov}\ (X, Y) = \mathcal{E}((X-\mu_X)(Y-\mu_Y)).$$

Ga na dat deze definitie zinvol is zodra $\mathcal{E}(X^2)$ en $\mathcal{E}(Y^2)$ bestaan.

Opgave: Toon aan dat $\mathrm{Cov}(X, Y) = \mathcal{E}(X \cdot Y) - \mu_X . \mu_Y$.

<u>Stelling 8.</u> Indien $X_1, X_2, \ldots, X_n$ stochastische funkties zijn met eindige spreidingen $\sigma_1, \sigma_2, \ldots, \sigma_n$ en $S_n = X_1 + X_2 + \ldots + X_n$ dan geldt

$$\mathrm{Var}\ (S_n) = \sum_{k=1}^{n} \sigma_k + 2 \sum_{i<j} \mathrm{Cov}\ (X_i, X_j)$$

Bewijs: $\mathrm{Var}\ (S_n) = \mathrm{Var}\ (X_1 + X_2 + \ldots + X_n) =$

$$= \mathcal{E}(X_1 + X_2 + \ldots + X_n - \mathcal{E}(X_1 + X_2 + \ldots + X_n))^2 =$$

$$= \mathcal{E}(X_1 + X_2 + \ldots + X_n - (\mu_1 + \mu_2 + \ldots + \mu_n))^2 =$$

$$= \mathcal{E}((X_1 - \mu_1) + (X_2 - \mu_2) + \ldots + (X_n - \mu_n))^2 =$$

$$= \mathcal{E}\left(\sum_{i=1}^{n} (X_i - \mu_i)^2 + 2 \sum_{i<j} (X_i - \mu_i)(X_j - \mu_j)\right) =$$

$$= \sum_{i=1}^{n} \sigma_i^2 + 2 \sum_{i<j} \mathrm{Cov}\ (X_i, X_j).$$

<u>Definitie.</u> Zijn X en Y stochastische funkties met eindige varianties $\neq 0$ dan wordt de <u>correlatie coëfficient</u> <u>van X en Y</u> (notatie $\rho(X, Y)$) gedefinieerd door

$$\rho(X, Y) = \mathrm{Cov}\,(X^*, Y^*)$$

Opgave: Toon aan dat $\rho(X, Y) = \dfrac{\mathrm{Cov}\,(X, Y)}{\sigma_X \cdot \sigma_Y}$.

Stelling 9. Altijd geldt $|\rho(X, Y)| \leq 1$.
Als $|\rho(X, Y)| = 1$ dan is $Y = a\,X + b$, waarbij a en b constanten zijn.

Bewijs: $\mathrm{Var}\,(X^* \pm Y^*) = \mathrm{Var}\,(X^*) \pm 2\,\mathrm{Cov}\,(X^*, Y^*) + \mathrm{Var}\,(Y^*) =$

$$= 2(1 \pm \rho(X, Y))$$

Aangezien $\mathrm{Var}\,(X^* \pm Y^*) \geq 0$, moet wel gelden

$$|\rho(X, Y)| \leq 1.$$

Als $\rho(X, Y) = 1$ dan geldt $\mathrm{Var}\,(X^* - Y^*) = 0$.
Maar dit houdt in dat $X^* - Y^*$ constant is (behalve misschien op de punten E_i waarvoor $p(E_i) = 0$). Op het "relevante deel" van E geldt dus $X^* - Y^* = c$ zodat $\dfrac{X-\mu_X}{\sigma_X} - \dfrac{Y-\mu_Y}{\sigma_Y} = c$ met als gevolg $Y = a\,X + b$ met $a = \dfrac{\sigma_2}{\sigma_1}$ en $b = \mu_Y - \dfrac{\sigma_2}{\sigma_1}\mu_X - c\,\sigma_2$. Het geval $\rho(X, Y)$ -1 kan analoog worden behandeld.

Stelling 10. (Ongelijkheid van Bienaymé - Čebyšev)
Zij X een stochastische funktie met verwachting μ en variantie σ^2.
Dan geldt voor elke $t \neq 0$ dat

$$P\{|X-\mu| \geq t\} \leq \frac{\sigma^2}{t^2}.$$

Hierbij is $P\{|X-\mu| \geq t\} = \sum_{|X(E_i)-\mu| \geq t} p(E_i)$.

Bewijs: $\sigma^2 = \text{Var}(X) = \sum_{E_i \in E} (X(E_i) - \mu)^2 p(E_i) \geqq$

$$\geqq \sum_{\substack{E_i \in E \\ |X(E_i)-\mu| \geqq t}} (X(E_i) - \mu)^2 p(E_i) \geqq t^2 \sum_{\substack{E_i \in E \\ |X(E_i)-\mu| \geqq t}} p(E_i) = t^2 \cdot P\{|X-\mu| \geqq t\}.$$

Als $t \neq 0$ dan geldt dus

$$P\{|X-\mu| \geqq t\} \leqq \frac{\sigma^2}{t^2}.$$

Opmerking. Voor $t = a\,\sigma\ (\neq 0)$ wordt dit

$$P\{|X-\mu| \geqq a \cdot \sigma\} \leqq \frac{1}{a^2}.$$

7. Stochastische onafhankelijkheid

Zij X een stochastische funktie op een WH-veld (E,p); dan definiëren we

$$P\{X = x\} \overset{\text{def}}{=} P\{X^{-1}(x)\}.$$

Zij ook Y een stochastische funktie op (E,p). Dan heten X en Y (stochastisch) onafhankelijk indien voor alle reële x en y geldt

$$P\{X^{-1}(x) \cap Y^{-1}(y)\} = P\{X^{-1}(x)\} \cdot P\{Y^{-1}(y)\}.$$

Voorbeeld. Zij $E = \{(0,0), (0,1), (1,0), (1,1)\}$ met $p(0,0) = \alpha \cdot \beta$, $p(0,1) = \alpha(1 - \beta)$, $p(1,0) = (1 - \alpha)\beta$ en $p(1,1) = (1 - \alpha)(1 - \beta)$ waarbij $0 < \alpha, \beta < 1$. Als voor elk punt $(x,y) \in E$ $X(x,y) = x$ en $Y(x,y) = y$, dan zijn de funkties X en Y stochastisch onafhankelijk. Ga dit na.

Zij X weer een stochastische funktie op het WH-veld (E,p). Aangezien E hoogstens aftelbaar veel punten E_i bevat zal X hoogstens aftelbaar veel verschillende waarden aannemen. De door X aangenomen waarden zullen we aangeven met $x_1, x_2, x_3, \ldots$. Definiëren we nu $f(x_i) = P\{X = x_i\}$ dan is de verzameling $\tilde{E} = \{x_1, x_2, x_3, \ldots\}$ met de daarop gedefinieerde funktie f een WH-veld. Dit is als volgt in te zien: het is duidelijk dat $f(x_i) \geqq 0$ voor $i = 1, 2, 3, \ldots$; daar we aangenomen hebben dat $x_i \neq x_j$ als $i \neq j$ kunnen we schrijven

$$1 = P\{E\} = P\{X^{-1}(x_1) \cup X^{-1}(x_2) \cup X^{-1}(x_3) \cup \ldots\} =$$

$$(X^{-1}(x_i) \text{ en } X^{-1}(x_j) \text{ zijn disjunct als } i \neq j)$$

$$= \sum_{i=1}^{\infty} P\{X^{-1}(x_i)\} = \sum_{i=1}^{\infty} P\{X = x_i\} = \sum_{i=1}^{\infty} f(x_i).$$

Stelling 11. Als $\mathcal{E}(X)$ bestaat dan is $\mathcal{E}(X) = \sum_{i=1}^{\infty} x_i f(x_i)$.

Bewijs. $\mathcal{E}(X) = \sum_{i=1}^{\infty} X(E_i)p(E_i) = \sum_{j=1}^{\infty} \sum_{X(E_i)=x_j} X(E_i)p(E_i) =$

$$= \sum_{j=1}^{\infty} x_j \sum_{X^{-1}(x_j)} p(E_i) = \sum_{j=1}^{\infty} x_j \, P\{X^{-1}(x_j)\} =$$

$$= \sum_{j=1}^{\infty} x_j \, f(x_j).$$

Ga na dat $\sum_{j=1}^{\infty} x_j \, f(x_j)$ inderdaad absoluut convergeert.

Stelling 12. Als X en Y (stochastisch) onafhankelijke stochastische funkties zijn op een WH-veld (E,p), terwijl Var(X) en Var(Y) bestaan, dan geldt dat $\mathcal{E}(X \cdot Y)$ bestaat en gelijk is aan

$$\mathcal{E}(X \cdot Y) = \mathcal{E}(X) \cdot \mathcal{E}(Y).$$

Bewijs. Dat $\mathcal{E}(X \cdot Y)$ bestaat is een gevolg van de ongelijkheid

$$|x \cdot y| \leq \frac{x^2 + y^2}{2} .$$

Verder is $\mathcal{E}(X \cdot Y) = \sum_{i=1}^{\infty} (X \cdot Y)(E_i) \cdot p(E_i) = \sum_{i=1}^{\infty} X(E_i)Y(E_i)p(E_i) =$

$$= \sum_{j,k=1}^{\infty} \sum_{\substack{X(E_i)=x_j \\ Y(E_i)=y_k}} X(E_i)Y(E_i)p(E_i) =$$

$$= \sum_{j,k=1}^{\infty} x_j \, y_k \sum_{\substack{X(E_i)=x_j \\ Y(E_i)=y_k}} p(E_i) =$$

$$= \sum_{j,k=1}^{\infty} x_j \, y_k \, P\{X^{-1}(x_j) \cap Y^{-1}(y_k)\} =$$

$$= \sum_{j,k=1}^{\infty} x_j \, y_k \, P\{X^{-1}(x_j)\} \cdot P\{Y^{-1}(y_k)\} =$$

$$= \sum_{j=1}^{\infty} x_j \, P\{X^{-1}(x_j)\} \cdot \sum_{k=1}^{\infty} y_k \, P\{Y^{-1}(y_k)\} =$$

$$= \mathcal{E}(X) \cdot \mathcal{E}(Y).$$

Stelling 13. Als X en Y stochastisch onafhankelijk zijn en beide eindige variantie bezitten dan geldt

$$\mathrm{Cov}(X,Y) = 0.$$

Bewijs. Dit volgt onmiddellijk uit de vorige stelling en de betrekking

$$\mathrm{Cov}(X,Y) = \mathcal{E}(X \cdot Y) - \mathcal{E}(X) \cdot \mathcal{E}(Y).$$

Stelling 14. Indien de op hetzelfde WH-veld (E,p) gedefinieerde stochastische funkties $X_1, X_2, \ldots, X_n$ paarsgewijs (stochastisch) onafhankelijk zijn, en allen eindige varianties bezitten, dan geldt dat

$$\mathrm{Var}(X_1 + X_2 + \ldots + X_n) \quad \text{bestaat en gelijk is}$$

aan

$$\mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_n).$$

Bewijs. Uit het bewijs van stelling 8 blijkt dat $\mathrm{Var}(X_1 + \ldots + X_n)$ bestaat en gelijk is aan

$$\sum_{i=1}^{n} \mathrm{Var}(X_i) + 2 \sum_{i<j} \mathrm{Cov}(X_i,X_j).$$

Daar X_i en X_j voor $i \neq j$ stochastisch onafhankelijk zijn is volgens de vorige stelling

$$\mathrm{Cov}(X_i,X_j) = 0 \quad (i \neq j).$$

Hieruit volgt de bewering onmiddellijk.

Opgaven. 1. Als X en Y stochastisch onafhankelijk zijn dan zijn ook de bij X en Y behorende genormaliseerde stochastische funkties X^* en Y^* onafhankelijk, en omgekeerd.

2. Zij $X_1, X_2, \ldots, X_n$ een stelsel paarsgewijs (stochastisch) onafhankelijke stochastische funkties op het WH-veld $E = \{0,1\}$ met $p(0) = \alpha$ en $p(1) = 1 - \alpha$ $(0 < \alpha < 1)$, zodanig dat

$$\left.\begin{array}{l} X_i(0) = 0 \\ \\ X_i(1) = 1 \end{array}\right\} \text{ voor alle } i = 1, 2, \ldots, n.$$

Indien $X = X_1 + X_2 + \ldots + X_n$, bereken dan μ_X en σ_X.
Bereken ook $\mathcal{E}(\frac{X}{n})$ en $\mathrm{Var}(\frac{X}{n})$.

Stelling 15. Zij $X_1, X_2, X_3, \ldots$ een oneindige rij paarsgewijs onafhankelijke stochastische funkties op het WH-veld (E,p) zodanig dat

$$\left.\begin{array}{l} \mathcal{E}(X_i) = \mu \\ \\ \mathrm{Var}(X_i) \leq G \end{array}\right\} \text{ voor } i = 1, 2, 3, \ldots$$

dan geldt voor elke positieve ε

$$\lim_{n\to\infty} P\left\{\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right\} = 0,$$

waarbij $S_n = X_1 + X_2 + \ldots + X_n$.

Bewijs. $\mathcal{E}(\frac{S_n}{n}) = \frac{1}{n}\mathcal{E}(S_n) = \frac{1}{n}\{\mathcal{E}(X_1) + \ldots + \mathcal{E}(X_n)\} = \frac{1}{n} \cdot n\mu = \mu.$

$$\mathrm{Var}(\frac{S_n}{n}) = \frac{1}{n^2} \cdot \mathrm{Var}(S_n) =$$

$$= \frac{1}{n^2}\{\mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n)\} \leq$$

$$\leq \frac{1}{n^2} \cdot nG = \frac{G}{n}.$$

Volgens de ongelijkheid van Bienaymé-Čebyšev met $X = \frac{S_n}{n}$ kunnen we dan schrijven

$$P\{|\frac{S_n}{n} - \mu| \geq \varepsilon\} \leq \frac{Var(\frac{S_n}{n})}{\varepsilon^2}$$

en wegens $Var(\frac{S_n}{n}) \leq \frac{G}{n}$ kunnen we concluderen dat

$$P\{|\frac{S_n}{n} - \mu| \geq \varepsilon\} \leq \frac{G}{n \cdot \varepsilon^2} .$$

Aangezien $\lim_{n \to \infty} \frac{G}{n \cdot \varepsilon^2} = 0$ volgt de stelling.

Colloquium "Waarschijnlijkheidsrekening" (1968-1969)

Spreker: H.G.J. Pijls

8. Stochastische onafhankelijkheid (vervolg).

Laat (E,p) een WH-veld zijn.

De gebeurtenissen A en B ($A, B \subset E$) heten onafhankelijk indien $P\{A \cap B\} = P\{A\}\, P\{B\}$.
Wij gaan nu definiëren wat we zullen verstaan onder: de gebeurtenissen A,B en C zijn onderling onafhankelijk.
Een rechtstreekse generalisatie van de eis in de bovenstaande definitie levert de conditie:

$$P\{A \cap B \cap C\} = P\{A\}\, P\{B\}\, P\{C\}. \qquad (*)$$

Deze voorwaarde (*) blijkt evenwel noch nodig noch voldoende voor de paarsgewijze onafhankelijkheid van A,B en C. Dit blijkt uit de volgende voorbeelden.

Voorbeeld 1.
Men werpt met een 4-kantige "dobbelsteen". We definiëren:
$A = \{(1),(4)\}$
$B = \{(2),(4)\}$
$C = \{(3),(4)\}$.
Dan zijn A,B en C paarsgewijs onafhankelijk maar er is niet voldaan aan (*); ga dit na.

Voorbeeld 2.
In een vaas bevinden zich 8 briefjes: $b_1, b_2, \ldots, b_8$. Het briefje b_k draagt als opschrift l_k ($k=1,2, \ldots,8$), waarbij:

$l_1 = (111)$ $\qquad$ $l_5 = (011)$
$l_2 = (101)$ $\qquad$ $l_6 = (010)$
$l_3 = (100)$ $\qquad$ $l_7 = (010)$
$l_4 = (100)$ $\qquad$ $l_8 = (001)$.

Men trekt een briefje uit deze vaas. Laat A_i (i=1,2,3) de gebeurtenis zijn dat het i-de getal op het getrokken briefje een 1 is. Dan geldt:

$$P\{A_1 \cap A_2 \cap A_3\} = P\{A_1\}\, P\{A_2\}\, P\{A_3\}.$$

Maar A_1, A_2 en A_3 zijn niet paarsgewijs onafhankelijk.

Wij zullen nu A,B en C onderling onafhankelijk noemen indien A,B en C paarsgewijs onafhankelijk zijn en bovendien (*) geldt.
De algemene definitie van onderlinge onafhankelijkheid van een n-tal gebeurtenissen luidt als volgt:

Definitie. De gebeurtenissen A_1, A_2, ..., A_n heten _onderling onafhankelijk_ indien voor alle combinaties $1 \leq i < j < k < \ldots \leq n$ geldt:

$$\begin{cases} P\{A_i \cap A_j\} = P\{A_i\}\, P\{A_j\} \\ P\{A_i \cap A_j \cap A_k\} = P\{A_i\}\, P\{A_j\}\, P\{A_k\} \\ \text{--------------------------} \\ \text{--------------------------} \\ P\{A_1 \cap A_2 \cap \ldots \cap A_n\} = P\{A_1\}\, P\{A_2\} \ldots P\{A_n\}. \end{cases}$$

(dit zijn 2^n-n-1 relaties).

Wij gaan nu over op stochastische funkties op (E,p). De stochastische funkties X en Y heten onafhankelijk indien de gebeurtenissen $X^{-1}(x)$ en $Y^{-1}(y)$ onafhankelijk zijn voor alle x en $y \in \mathbb{R}$.
Als X_1, X_2, ..., X_n stochastische funkties zijn, dan definiëren we:

$$P\{X_1=a_1,\, X_2=a_2,\, \ldots,\, X_n=a_n\} \stackrel{\text{def}}{=} P\{X_1^{-1}(a_1) \cap X_2^{-1}(a_2) \cap \ldots \cap X_n^{-1}(a_n)\}.$$

Definitie. De stochastische funkties X_1, X_2, ..., X_n heten _onderling onafhankelijk_, indien de gebeurtenissen $X_1^{-1}(a_1)$, $X_2^{-1}(a_2)$, ..., $X_n^{-1}(a_n)$ onderling onafhankelijk zijn voor elk n-tal reële getallen a_1, a_2, ...,a_n.

We bewijzen nu dat stochastische funkties onderling onafhankelijk zijn, indien voldaan is aan een relatie die correspondeert met de betrekking (*).

Stelling 16. De stochastische funkties $X_1, X_2, \ldots, X_n$ zijn onderling onafhankelijk, indien voor elk n-tal reële getallen $a_1, a_2, \ldots, a_n$ geldt:

$$P\{X_1=a_1, X_2=a_2, \ldots, X_n=a_n\} = P\{X_1=a_1\} P\{X_2=a_2\} \ldots P\{X_n=a_n\}.$$

Bewijs.
We bewijzen het geval n=3. Het algemene geval wordt analoog bewezen. Op grond van symmetrie-overwegingen behoeven we slechts te bewijzen dat voor elk paar reële getallen a en b geldt:

$$P\{X_1=a, X_2=b\} = P\{X_1=a\} P\{X_2=b\}.$$

Laat $\{c_1, c_2, \ldots\}$ de waardenverzameling van X_3 zijn. Dan geldt:

$$\begin{aligned} P\{X_1=a, X_2=b\} &= P\{\bigcup_{k=1}^{\infty} (X_1^{-1}(a) \cap X_2^{-1}(b) \cap X_3^{-1}(c_k))\} \\ &= \sum_{k=1}^{\infty} P\{X_1^{-1}(a) \cap X_2^{-1}(b) \cap X_3^{-1}(c_k)\} \\ &= \sum_{k=1}^{\infty} P\{X_1=a\}.\ P\{X_2=b\}.\ P\{X_3^{-1}(c_k)\} \\ &= P\{X_1=a\}.\ P\{X_2=b\}.\ P\{\bigcup_{k=1}^{\infty} X_3^{-1}(c_k)\} \\ &= P\{X_1=a\}.\ P\{X_2=b\}. \end{aligned}$$

Bij het bewijs van de volgende stelling zullen we de volgende lemma's gebruiken:

Lemma 1. Laat $\mathfrak{A} = (A_k)_{k=1}^{\infty}$ een stelsel van elkaar uitsluitende gebeurtenissen zijn; zo ook het stelsel $\mathfrak{B} = (B_k)_{k=1}^{\infty}$. Laten A en B onafhankelijk zijn voor alle $A \in \mathfrak{A}$, $B \in \mathfrak{B}$. Dan zijn ook $\bigcup_{k=1}^{\infty} A_k$ en $\bigcup_{l=1}^{\infty} B_l$ onafhankelijk.

Bewijs.

Dit volgt uit de volgende gelijkheden.

$$P\{(\bigcup_{k=1}^{\infty} A_k) \cap (\bigcup_{l=1}^{\infty} B_l)\} = P\{\bigcup_{k,l=1}^{\infty} (A_k \cap B_l)\} =$$

$$= \sum_{k,l=1}^{\infty} P\{A_k \cap B_l\}$$

$$= \sum_{k,l=1}^{\infty} P\{A_k\} \cdot P\{B_l\}$$

$$= P\{\bigcup_{k=1}^{\infty} A_k\} \cdot P\{\bigcup_{l=1}^{\infty} B_l\}.$$

Lemma 2. Als de gebeurtenissen A,B en C onderling onafhankelijk zijn, dan zijn ook $A \cap B$ en C onafhankelijk.

Bewijs.

$$P\{(A \cap B) \cap C\} = P\{A\}\ P\{B\}\ P\{C\}$$

$$= P\{A \cap B\}\ P\{C\}.$$

<u>Stelling 17.</u> Laten $X_1, X_2, \ldots, X_n$ onderling onafhankelijke stochastische funkties zijn op een WH-veld (E,p). Als U een funktie is van $X_1, X_2, \ldots, X_k$ en V een funktie van $X_{k+1}, \ldots, X_n$, waarbij $k < n$, dan zijn de stochastische funkties U en V onafhankelijk.

Bewijs.

We bewijzen het geval n=3. Het bewijs van het algemene geval verloopt analoog.

Op grond van symmetrie-overwegingen behoeven we slechts het volgende te bewijzen:

Laten X_1, X_2, X_3 onderling onafhankelijke stochastische funkties zijn. Als $f: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ en $g: \mathbb{R} \to \mathbb{R}$ willekeurige funkties zijn, dan zijn ook de stochastische funkties $Y_1 = f(X_1,X_2)$ en $Y_2 = g(X_3)$ onafhankelijk.

We moeten daarvoor aantonen dat de gebeurtenissen $Y_1^{-1}(y_1)$ en $Y_2^{-1}(y_2)$ onafhankelijk zijn voor alle $y_1, y_2 \in \mathbb{R}$.

Nu geldt:

$$Y_1^{-1}(y_1) = \bigcup_{(x_1,x_2)\in f^{-1}(y_1)} (X_1^{-1}(x_1) \cap X_2^{-1}(x_2)).$$

Immers, voor $E_i \in E$ geldt:

$Y_1(E_i) = y_1$ d.e.s.d. als $\exists x_1,x_2 \in \mathbb{R}$ met $X_1(E_i) = x_1$

$X_2(E_i) = x_2$

en $f(x_1,x_2) = y_1$.

Evenzo geldt:

$$Y_2^{-1}(y_2) = \bigcup_{x_3\in g^{-1}(y_2)} X_3^{-1}(x_3).$$

Nu is $X_1^{-1}(x_1) \cap X_2^{-1}(x_2) \neq \emptyset$ voor slechts aftelbaar veel paren (x_1,x_2) (immers, de waardenverzamelingen van X_1 en X_2 zijn beide aftelbaar). Verder is het stelsel $\{X_1^{-1}(x_1) \cap X_2^{-1}(x_2) \mid (x_1,x_2)\in f^{-1}(y_1)\}$ een stelsel van elkaar uitsluitende gebeurtenissen.
Evenzo is $X_3^{-1}(x_3) \neq \emptyset$ voor slechts aftelbaar veel waarden van x_3 en is het stelsel $\{X_3^{-1}(x_3) \mid x_3 \in g^{-1}(y_2)\}$ een stelsel van elkaar uitsluitende gebeurtenissen.

Uit het gegeven, dat X_1,X_2 en X_3 onderling onafhankelijk zijn, volgt, dat de gebeurtenissen $X_1^{-1}(x_1)\cap X_2^{-1}(x_2)$ en $X_3^{-1}(x_3)$ onafhankelijk zijn (Lemma 2).

Uit Lemma 1 volgt dan dat $Y_1^{-1}(y_1)$ en Y_2^{-1} onafhankelijk zijn. Hiermee is het geval n=3 bewezen.

<u>Opgaven</u>

1. Laten A,B en C onderling onafhankelijke gebeurtenissen zijn. Dan zijn $A \cup B$ en C ook onafhankelijk.

2. Laten $X_1,X_2, \ldots, X_n$ onderling onafhankelijke stochastische funkties zijn. Dan geldt:

$$\mathcal{E}(X_1 . X_2 \ldots X_n) = \mathcal{E}(X_1).\mathcal{E}(X_2) \ldots \mathcal{E}(X_n).$$

9. Genererende funkties.

Definitie. Laat X een stochastische funktie zijn op een WH-veld (E,p) met waarden $x_1, x_2, x_3, \ldots$. Op de verzameling $W = \{x_1, x_2, \ldots\}$ definiëren we de funktie f als volgt:

$$f(x_i) \overset{\text{def}}{=} P\{X = x_i\}$$

Dan is (W,f) weer een WH-veld, dat de verdeling van X heet.
Als de waarden van de stochastische funktie X gehele getallen ≥ 0 zijn (i.e. $W \subset \mathbb{N} \cup \{0\}$), dan kunnen we de verdeling van X identificeren met de rij $\{a_n\}_{n=0}^{\infty}$, waarbij:

$$a_n = P\{X = n\}.$$

De machtreeks

$$P(s) = \sum_{n=0}^{\infty} a_n s^n \qquad (*)$$

heet de genererende funktie van de stochastische functie X (of: de genererende funktie van de rij $\{a_n\}$).
Als R de convergentiestraal van de machtreeks (*) voorstelt dan is $R \geq 1$ (immers voor s=1 convergeert de reeks). De genererende funktie P(s) van de verdeling $\{a_n\}_{n=0}^{\infty}$ is dus een analytische funktie van s voor $|s| < 1$.

In verband met bovenstaande definitie beschouwen we in deze paragraaf slechts stochastische funkties met waarden in $\mathbb{N} \cup \{0\}$.

Uit bovenstaande definitie volgt dat de genererende funktie van een stochastische funktie X bepaald wordt door de verdeling van X. Omgekeerd geldt: als P(s) de genererende funktie van een stochastische funktie X is (P(s) is dan noodzakelijk analytisch in s=0), dan kunnen we de verdeling van X berekenen (ontwikkel P(s) bij s=0 in een machtreeks).

Voorbeeld.
Men werpt met een dobbelsteen. Laat X het aantal ogen zijn dat men werpt. De genererende functie P(s) van X is:

$$P(s) = \frac{1}{6}(s + s^2 + s^3 + s^4 + s^5 + s^6)$$

$$= \frac{s(1-s^6)}{6(1-s)} .$$

Laten nu gegeven zijn twee stochastische functies X en Y op een zeker WH-veld (E,p), die stochastisch onafhankelijk zijn. Laat de verdeling van X resp. Y gelijk zijn aan $\{a_n\}$ resp. $\{b_n\}$.
Laat $\{c_n\}$ de verdeling zijn van de stochastische functie S=X+Y.
Dan is

$$c_n = P\{S = n\} = P\{X + Y = n\}.$$

De gebeurtenis (X+Y=n) is de vereniging van de volgende gebeurtenissen:

$$(X=0,\ Y=n),\ (X=1,\ Y=n-1),\ \dots,\ (X=k,\ Y=n-k),\ \dots,\ (X=n,\ Y=0).$$

Dus: $$P\{X+Y=n\} = P\{\bigcup_{k=0}^{n} (X=k,\ Y=n-k)\}.$$

Daar de genoemde gebeurtenissen elkaar uitsluiten geldt:

$$P\{X+Y=n\} = \sum_{k=0}^{n} P\{X=k,\ Y=n-k\}.$$

X en Y zijn stochastisch onafhankelijk en dus:

$$P\{X+Y=n\} = \sum_{k=0}^{n} P\{X=k\}\ P\{Y=n-k\}.$$

Ofwel: $$c_n = \sum_{k=0}^{n} a_k\ b_{n-k}.$$

We voeren nu de volgende definitie in:

<u>Definitie</u> Laten $\{a_n\}$ en $\{b_n\}$ rijen van reële getallen zijn. De rij $\{c_n\}$, gedefinieerd door:

$$c_n = \sum_{k=0}^{n} a_k\ b_{n-k}$$

heet de <u>convolutie</u> van $\{a_n\}$ en $\{b_n\}$.
Notatie: $\{c_n\} = \{a_n\} * \{b_n\}$.

We hebben dus bewezen dat de verdeling van X+Y gelijk is aan $\{a_n\} * \{b_n\}$.

We vragen nu naar de genererende funktie van $X + Y$.

Lemma. Laten $\{a_n\}$ en $\{b_n\}$ rijen van reële getallen zijn zodanig dat de machtreeksen

$$P(s) = \sum_{n=0}^{\infty} a_n s^n \text{ en } Q(s) = \sum_{n=0}^{\infty} b_n s^n$$

beide een convergentiestraal ≥ 1 hebben. Laat de rij $\{c_n\}$ gedefinieerd zijn door

$$\{c_n\} = \{a_n\} * \{b_n\}.$$

Dan heeft de machtreeks

$$R(s) = \sum_{n=0}^{\infty} c_n s^n$$

ook een convergentiestraal ≥ 1 en bovendien geldt

$$R(s) = P(s)\, Q(s).$$

Anders geformuleerd:
Als $P(s)$ resp. $Q(s)$ de genererende funktie van $\{a_n\}$ resp. $\{b_n\}$ is, dan is $R(s) = P(s)\, Q(s)$ de genererende funktie van $\{a_n\} * \{b_n\}$.

We vatten deze resultaten samen in de volgende stelling.

Stelling 18. Laten X en Y onafhankelijke stochastische funkties zijn met waarden in $\mathbb{N} \cup \{0\}$.
Laat de verdeling van X resp. Y gelijk zijn $\{a_n\}$ resp. $\{b_n\}$ en laat de genererende funktie van X resp. Y gelijk zijn aan $P(s)$ resp. $Q(s)$.
Dan geldt:
a) De stochastische funktie $X + Y$ heeft de verdeling $\{a_n\} * \{b_n\}$.
b) De stochastische funktie $X + Y$ heeft $P(s)\, Q(s)$ tot genererende funktie.

Gevolg. $X_1, X_2, \ldots, X_n$ zijn onderling onafhankelijke stochastische funkties met waarden in $\mathbb{N} \cup \{0\}$. Laat de genererende funktie van X_k gelijk zijn aan $P_k(s)$ $(k=1,2, \ldots,n)$. De stochastische funktie $S = X_1 + X_2 + \ldots + X_n$ heeft dan $P_1(s) . P_2(s) \ldots P_n(s)$ tot genererende funktie.

Bewijs

Dit gevolg wordt bewezen door Stelling 18 toe te passen op de funkties X_{k+1} en $X_1 + X_2 + \ldots + X_k$ voor $k=1,2, \ldots,n-1$; dat X_{k+1} en $X_1 + X_2 + \ldots + X_k$ onafhankelijk zijn volgt uit Stelling 17.

Opgaven.

1. X is een stochastische funktie gedefinieerd op zeker WH-veld met waarden in $\mathbb{N} \cup \{0\}$; $\mathcal{E}(X)$ en $\mathcal{E}(X^2)$ bestaan. Laat de genererende funktie van X gelijk zijn aan P(s).

   Bewijs:

   a) $\mathcal{E}(X) = P'(1)$

   b) $\mathcal{E}(X^2) = P'(1) + P''(1)$

   c) $\mathrm{Var}\ (X) = P'(1) - P'^2(1) + P''(1)$.

2. a) X is een stochastische funktie met waarden in $\mathbb{N} \cup \{0\}$. De genererende funktie van X is gelijk aan P(s).

   Bewijs:

   $P(s) = \mathcal{E}(s^X) \quad (0 \leq s \leq 1)$.

   b) Bewijs Stelling 18b) m.b.v. a) en Stelling 17.

3. a) Laten X en Y onafhankelijke stochastische funkties zijn, gedefinieerd op een WH-veld (E,p).

   De verdeling van X resp. Y is gelijk aan (W_1,f_1) resp. (W_2,f_2).

   Zij $W = W_1 \cap W_2$.

   Dan geldt:

   $$P\{X = Y\} = \sum_{x \in W} f(x)g(x)$$

   b) Zij $\mathcal{F}$ de verzameling van alle stochastische funkties gedefinieerd op (E,p). Zij $X \in \mathcal{F}$. Dan geldt: X en Y zijn onafhankelijk voor alle $Y \in \mathcal{F}$ d.e.s.d. als X constant is.

4. Men werpt n maal met een "zuivere N-kantige dobbelsteen". Zij X_k het aantal ogen dat men bij de k-de worp werpt. En zij S het totale aantal geworpen ogen.

   a) Bewijs dat $X_1,X_2, \ldots,X_n$ onderling onafhankelijk zijn.

   b) Laat zien hoe men $P\{S = k\}$ berekent.

5. (Bernoulli-proefneming).

Een zeker experiment heeft slechts twee mogelijke uitkomsten: S(succes) en M(mislukking). De kans op S is gelijk aan p. Men voert dit experiment n maal uit. Bereken de kans op precies k successen.

6. (Poisson-proefnemingen).

Men voert n experimenten uit. Ieder experiment heeft de uitkomstmogelijkheden: S(succes) en M(mislukking). Het k-de experiment heeft een kans p_k op succes. Bereken de kans op precies k successen.

Vraagstukken Waarschijnlijkheidsrekening (1968-1969)

1. Men heeft n verschillende paren schoenen in een kast. Als men uit deze kast aselect m schoenen kiest hoe groot is dan de kans dat er zich onder deze m schoenen minstens één paar bevindt?

2. Men werpt 5 maal met een zuivere munt. Hoe groot is de kans dat men minstens drie maal achtereen "kruis" werpt?

3. Men heeft een vaas met daarin n briefjes, genummerd met 1, 2, 3, ..., n. Uit de vaas worden aselect (zonder teruglegging) r briefjes gekozen. Hoe groot is de kans dat onder de gekozen briefjes de nummers 1, 2, 3, ..., N voorkomen.

4. Men werpt drie maal met een zuivere dobbelsteen. Als bekend is dat er tenminste één maal 6 geworpen is hoe groot is dan de kans dat er twee of meer 6-en geworpen zijn?

6. Hoe groot is de waarschijnlijkheid, dat in een pak van 52 bridgekaarten (na goed schudden) geen van de 4 azen op een andere aas ligt?

7. (Bridge).
Noord en Zuid hebben samen 10 troeven, waaronder de Aas. Hoe groot is de kans dat de overige 3 troeven in één hand zijn. Als verder bekend is, dat zich onder deze 3 troeven de Heer bevindt, hoe groot is dan de kans dat deze "sec" zit?

8. Een open snoer met n kralen wordt direct links van een aselect gekozen kraal in tweeën geknipt. Met de rechterhelft doet men hetzelfde. Bepaal de verdeling van het aantal kralen in het middenstuk.

9. Een zeker gebouw heeft n verdiepingen. In de lift bevinden zich r personen. Hoe groot is de kans dat de lift precies k verdiepingen moet aandoen?

10. (Het lucifersdoosjes-probleem van Banach).
Een zekere wiskundige heeft twee doosjes lucifers op zak (ieder doosje bevat N lucifers). Steeds als hij nu een lucifer nodig heeft, kiest hij aselect een doosje en neemt daaruit een lucifer. Op zeker moment ontdekt hij, voor het eerst, dat het gekozen doosje leeg is. Hoe groot is de kans dat het andere doosje dan nog precies k lucifers bevat?
Hoe groot is de verwachting van het aantal lucifers in dit doosje?
Bereken deze verwachting voor $N = 50$.

11. Op het WH-veld $E = \prod_{i=1}^{n} \{0, 1\}$ met $p(a_1, a_2, \ldots, a_n) = \alpha^i(1-\alpha)^{n-i}$, $(0 < \alpha < 1)$, waarbij i = het aantal nullen in $\{a_1, a_2, \ldots, a_n\}$, beschouwt men de stochastische funkties $X_1, X_2, \ldots, X_n$ die als volgt gedefinieerd zijn

$$X_i(a_1, a_2, \ldots, a_n) = 0 \text{ als } a_i = 0,$$
$$X_i(a_1, a_2, \ldots, a_n) = 1 \text{ als } a_i = 1.$$

Toon aan dat X_i en X_j $(i \neq j)$ stochastisch onafhankelijk zijn.
Indien $X = X_1 + X_2 + \ldots + X_n$, bereken dan
μ_X, σ_X , $\mathcal{E}(\frac{X}{n})$ en Var $(\frac{X}{n})$.

12. Zij X een stochastische funktie op het WH-veld (E, p) en $\psi : \mathbb{R} \to \mathbb{R}$. Dan is ook $\psi \circ X$ met $(\psi \circ X)(E_i) = \psi(X(E_i))$ een stochastische funktie op (E, p). Als $\mathcal{E}(\psi \circ X)$ bestaat dan geldt

$$\mathcal{E}(\psi \circ X) = \sum_k \psi(x_k)\, f(x_k).$$

Toon dit aan.

13. Indien $\psi: \mathbb{R} \to \mathbb{R}$ niet negatief, even en op $x \geq 0$ monotoon stijgend is en X is een stochastische funktie op (E,p) dan is, onder de voorwaarde dat $\mathcal{E}(\psi \circ X)$ bestaat,

$$P\{|X| \geq t\} \leq \frac{\mathcal{E}(\psi \circ X)}{\psi(t)} \quad (t > 0).$$

14. Een stad heeft n woonblokken waarvan er n_j zijn met x_j inwoners ($n = n_1 + n_2 + \ldots$). Uit deze n blokken kiest men aselect r verschillende en telt hierin het aantal inwoners. Zij $X_1, X_2, \ldots, X_r$ het aantal inwoners in de r gekozen blokken en zij $S_r = X_1 + X_2 + \ldots + X_r$.

Bewijs dat $\mathcal{E}(X_i) = \mathcal{E}(X_j)$ en $\mathrm{Var}(X_i) = \mathrm{Var}(X_j)$.

Indien $\mathcal{E}(X_i) = m$ en $\mathrm{Var}(X_i) = a$, druk dan $\mathcal{E}(S_r)$ en $\mathrm{Var}(S_r)$ uit in m en a.

15. Zij S_n het aantal successen in een reeks van n onafhankelijke steekproeven met kansen resp. $p_1, p_2, \ldots, p_n$ op succes.

De gemiddelde kans op succes is $a_n = \dfrac{p_1 + \ldots + p_n}{n}$.

Bewijs dat bij vaste gegeven a_n de variantie van S_n maximaal is als alle p_i gelijk zijn.

Colloquium "Waarschijnlijkheidsrekening" (1968-1969)

Spreker: A. Hordijk

9. Bernoulli wandelingen

Veronderstel dat Peter en Paul het volgende spel spelen: Peter gooit met een munt; als de worp resulteert in kruis dan geeft hij Paul een gulden en als de worp resulteert in munt dan krijgt hij een gulden van Paul. Peter gooit weer en weer geeft of krijgt hij een gulden al naar de munt valt. We veronderstellen dat Peter en Paul onvermoeibare spelers zijn, zodat ze in dit spel oneindig vaak met de munt werpen.
Laten we verder veronderstellen dat ze met een munt gooien waarvoor de kans op munt p $(0 < p < 1)$ en de kans op kruis q is $(p + q = 1)$.

We willen voor dit spel een kansrekenkundig model gaan opbouwen, of anders gezegd, we willen een WH-veld construeren dat aan dit spel beantwoordt.

De elementaire gebeurtenissen van dit WH-veld zullen zijn de oneindige rijen van plus en min énen, waarbij we afspreken dat als op de k-de plaats in de rij een +1 (resp. -1) staat, de k-de worp in munt (resp. kruis) resulteerde.
Als Ω_0 de verzameling is van alle elementaire gebeurtenissen en als we een element uit Ω_0 met ω aanduiden, dan is dus:

$$\Omega_0 = \{\omega \mid \mathbb{N} \overset{\omega}{\to} \Delta\} \qquad \text{waarin } \Delta = \{+1,-1\}.$$

Uit de verzamelingenleer weten we dat Ω_0 meer dan aftelbaar veel elementen bevat. Willen we Ω_0 tot een WH-veld maken dan zullen we anders te werk moeten gaan dan bij de tot nu toe beschouwde verzamelingen E waarin hoogstens aftelbaar veel elementaire gebeurtenissen waren. Immers, in het geval van hoogstens aftelbaar veel elementaire gebeurtenissen kunnen we, uitgaande van de kansen op de elementaire gebeurtenissen, de kans op iedere deelverzameling vinden door de kansen op de elementaire gebeurtenissen in die deelverzameling op te tellen. Dit lukt niet meer bij Ω_0.

Daarom gaan we eerst definiëren wat we in het meest algemene geval onder een WH-veld zullen verstaan.

Zij Ω een totaal willekeurige verzameling en zij $\mathcal{A}$ een klasse van deelverzamelingen van Ω die we de <u>meetbare verzamelingen</u> of ook wel de gebeurtenissen zullen noemen. Zoals het woord meetbaar al suggereert moet voor iedere verzameling $A \in \mathcal{A}$, de kans $P(A)$ op de gebeurtenis A gegeven zijn.

Ω is een verzameling waarvan we in ieder geval willen dat hij meetbaar is. Voorts willen we dat de kans op Ω, oftewel de kans op de zekere gebeurtenis, gelijk aan 1 is.

i) $\quad \Omega \in \mathcal{A}$ en $P(\Omega) = 1$

Als A een gebeurtenis is, dan ligt het voor de hand om te eisen dat aan het niet optreden van A ook een kans wordt toegekend en wel zodanig dat

ii) $\quad A \in \mathcal{A} \Rightarrow \neg A \in \mathcal{A}$ en $P(A) + P(\neg A) = 1$

Als $\{A_i\}_{i=1}^{\infty}$ een rij van gebeurtenissen is dan eisen we dat het optreden van één of meer van de gebeurtenissen A_i ook een gebeurtenis is. Als bovendien de gebeurtenissen A_i elkaar uitsluiten dan willen we dat de kans op het optreden van minstens één van de gebeurtenissen A_i gelijk is aan de som van de kansen op de gebeurtenissen A_i.

iii) $\quad \{A_i\}_{i=1}^{\infty} \subset \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ en

$$A_i \cap A_j = \emptyset \text{ voor } i \neq j \Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

<u>Definitie</u>

Zij Ω een verzameling, $\mathcal{A}$ een klasse van deelverzamelingen van Ω en P een funktie die aan iedere A behorend tot de klasse $\mathcal{A}$ een getal $P(A)$ met $0 \leq P(A) \leq 1$ toekent. Dan vormt het drietal $(\Omega, \mathcal{A}, P)$ een WH-veld indien aan de eisen i), ii) en iii) voldaan wordt.

Opmerking

De klasse van deelverzamelingen $\mathcal{A}$ wordt een σ-algebra en de funktie P wordt een kansmaat genoemd.

Ga na dat in het geval Ω hoogstens aftelbaar veel punten bevat en de één-puntsverzamelingen tot de klasse $\mathcal{A}$ behoren, het WH-veld $(\Omega,\mathcal{A},P)$ voldoet aan de definitie van een discreet WH-veld zoals dit gegeven is op pagina 3.

Alvorens verder te gaan willen we nog enkele eigenschappen van een WH-veld $(\Omega,\mathcal{A},P)$ afleiden.

1. $A,B\in\mathcal{A} \implies A\cup B\in\mathcal{A}$

$$\text{immers: } \Omega\in\mathcal{A} \overset{\text{ii)}}{\implies} \neg\Omega = \emptyset\in\mathcal{A}$$
$$\overset{\text{iii)}}{\implies} A\cup B\cup\emptyset\cup\emptyset\cup\ldots = A\cup B\in\mathcal{A}$$

2. $A,B\in\mathcal{A} \implies A\cap B\in\mathcal{A}$

$$\text{immers: } A,B\in\mathcal{A} \overset{\text{ii) en 1.}}{\implies} \neg A\cup\neg B = \neg(A\cap B)\in\mathcal{A}$$
$$\overset{\text{ii)}}{\implies} A\cap B\in\mathcal{A}$$

Geheel analoog aan 1. en 2. kunnen we bewijzen dat iedere vereniging van eindig veel verzamelingen uit $\mathcal{A}$ weer een verzameling uit $\mathcal{A}$ is, en dat iedere doorsnede van eindig veel, ja zelfs dat iedere doorsnede van aftelbaar veel verzamelingen uit $\mathcal{A}$ weer een verzameling behorend tot $\mathcal{A}$ is.

Nu we gedefiniëerd hebben wat we in het algemeen onder een WH-veld zullen verstaan, zullen we nagaan waaraan het WH-veld, noem het maar $(\Omega_0, \mathcal{A}_0, P_0)$, moet voldoen wil het als model kunnen dienen voor het spel tussen Peter en Paul.

Daar we Ω_0 reeds gedefiniëerd hebben, rest ons nog $\mathcal{A}_0$ en P_0 te definiëren. We willen dat de kans dat de k-de worp in munt resp. kruis resulteert p resp. q is. Laten we de "k-de coördinaat" van ω, dat is dus de uitkomst van de k-de worp, voorstellen door $\chi_k(\omega)$.

a) $\forall k \in \mathbb{N}$ geldt: $\{\chi_k(\omega) = \pm 1\} \in \mathcal{A}_0$ en $P_0\{\chi_k(\omega) = \pm 1\} = \{^p_q$.

We willen dat de kans dat b.v. de k-de worp in munt en de l-de worp ($k \neq l$) in kruis resulteert gelijk is aan pq.

b) $\forall k,l \in \mathbb{N}$, $k \neq l$ en $\forall \delta_1, \delta_2 \in \Delta$ geldt:

$$\{\chi_k(\omega) = \delta_1, \chi_l(\omega) = \delta_2\} = \{\chi_k(\omega) = \delta_1\} \cap \{\chi_l(\omega) = \delta_2\} \in \mathcal{A}_0$$

en

$$P_0\{\chi_k(\omega) = \delta_1, \chi_l(\omega) = \delta_2\} = P_0\{\chi_k(\omega) = \delta_1\}.P_0\{\chi_l(\omega) = \delta_2\}$$

We willen de eis b) verscherpen in die zin dat we een uitspraak analoog aan b) willen doen voor iedere eindige combinatie van worpen.

c) $\forall m \in \mathbb{N}$, $\forall k_1, k_2, \ldots, k_m \in \mathbb{N}$, $k_i \neq k_j$ voor $i \neq j$ en $\forall \delta_1, \delta_2, \ldots, \delta_m \in \Delta$

geldt:

$$\{\chi_{k_1}(\omega) = \delta_1, \chi_{k_2}(\omega) = \delta_2, \ldots, \chi_{k_m}(\omega) = \delta_m\} = \{\chi_{k_1}(\omega) = \delta_1\} \cap \{\chi_{k_2}(\omega) = \delta_2\} \cap \ldots \cap \{\chi_{k_m}(\omega) = \delta_m\} \in \mathcal{A}_0$$

en

$$P_0\{\chi_{k_1}(\omega) = \delta_1, \chi_{k_2}(\omega) = \delta_2, \ldots, \chi_{k_m}(\omega) = \delta_m\} = P_0\{\chi_{k_1}(\omega) = \delta_1\}.P_0\{\chi_{k_2}(\omega) = \delta_2\}. \ldots . P_0\{\chi_{k_m}(\omega) = \delta_m\}$$

Nu blijken er meerdere σ-algebra's in Ω_0 te zijn die voldoen aan de eisen a), b) en c). We leggen $\mathcal{A}_0$ vast door voor $\mathcal{A}_0$ te nemen de kleinst mogelijke klasse van deelverzamelingen van Ω_0 die én een σ-algebra is én voldoet aan a), b) en c). Dan is $\mathcal{A}_0$ precies de doorsnede van alle σ-algebra's die voldoen aan a), b) en c). Met betrekking tot de kansmaat P_0 lenen we een stelling uit de maattheorie die we hier niet zullen bewijzen.

Stelling Er is precies één kansmaat P_0 die gedefiniëerd is op de klasse $\mathcal{A}_0$ en voldoet aan a), b) en c).

Opmerkingen: 1. De uitdrukking dat de worpen onderling onafhankelijk zijn vinden we als een direkt gevolg van b) in ons model terug als: de stochastische funkties χ_k en χ_l zijn voor $k \neq l$ stochastisch onafhankelijk.

2. Een deelverzameling van Ω_0 die we verkrijgen door aan eindig veel coördinaten van ω, zeg $k_1, k_2, \ldots, k_m$, voorwaarden op te leggen is de vereniging van een eindig aantal elkaar uitsluitende verzamelingen van het type dat in voorwaarde c) is beschreven. Dus is zo'n verzameling meetbaar en de kans erop is de som van de kansen op die "c)"-verzamelingen. We noemen zo'n verzameling "een gebeurtenis die alleen afhangt van de worpen $k_1, k_2, \ldots, k_m$". Als illustratie:

$$\{\chi_1(\omega)<2, \chi_1(\omega)+\chi_2(\omega)<2, \chi_1(\omega)+\chi_2(\omega)+\chi_3(\omega)<2, \chi_1(\omega)+\chi_2(\omega)+\chi_3(\omega)+\chi_4(\omega)=2\} = \{\chi_1(\omega)=+1, \chi_2(\omega)=-1, \chi_3(\omega)=+1, \chi_4(\omega)=+1\} \cup \{\chi_1(\omega)=-1, \chi_2(\omega)=+1, \chi_3(\omega)=+1, \chi_4(\omega)=+1\}$$

en

$$P_0\{\chi_1(\omega)<2, \chi_1(\omega)+\chi_2(\omega)<2, \chi_1(\omega)+\chi_2(\omega)+\chi_3(\omega)<2, \chi_1(\omega)+\chi_2(\omega)+\chi_3(\omega)+\chi_4(\omega)=2\} = P_0\{\chi_1(\omega)=+1, \chi_2(\omega)=-1, \chi_3(\omega)=+1, \chi_4(\omega)=+1\} + P_0\{\chi_1(\omega)=-1, \chi_2(\omega)=+1, \chi_3(\omega)=+1, \chi_4(\omega)=+1\} = pqpp + qppp = 2p^3q.$$

3. Omdat we in deze paragraaf steeds de kansmaat P_0 zullen gebruiken zullen we voor P_0 in het vervolg P schrijven.

Naast de stochastische funkties χ_k introduceren we sommen van deze stochastische funkties. Zij $S_n = \chi_1 + \chi_2 + \ldots + \chi_n$; S_n is de winst of het verlies dat Peter heeft na precies n worpen.

Het viertal $(\Omega_0, \mathcal{Q}_0, P, \{S_n\}_{n=1}^{\infty})$ noemen we een Bernoulli wandeling.

Voor ieder positief of negatief geheel getal k definiëren we een stochastische funktie $T^{(k)}$ zodanig dat

$$T^{(k)}(\omega) = \min\{m \mid S_m(\omega) = k\}^{*}.$$

We noemen $T^{(k)}$ "de wachttijd tot het bereiken van niveau k". $\{T^{(k)}(\omega) = n\}$ is de verzameling van die ω's waarvoor we na precies n stappen voor het eerst niveau k bereiken. Het is een gebeurtenis die bepaald wordt door de eerste n worpen en is dus een meetbare verzameling. Zij $\lambda_n^{(k)} = P\{T^{(k)}(\omega)=n\}$. Laat verder $L^{(k)}(s) = \sum_{n=1}^{\infty} \lambda_n^{(k)} s^n$ de genererende funktie zijn van de rij $\{\lambda_n^{(k)}\}_{n=1}^{\infty}$.

---

*We spreken af dat als er voor ω geen enkele waarde van m bestaat zodanig dat $S_m(\omega) = k$, we $T^{(k)}(\omega) = \infty$ stellen.

Bewering: de machtreeks $L^{(k)}(s)$ convergeert voor $-1\leq s\leq +1$. Daar de gebeurtenissen $\{T^{(k)}(\omega) = n\}$ en $\{T^{(k)}(\omega) = m\}$ elkaar voor $n\neq m$ uitsluiten geldt:

$$\sum_{n=1}^{\infty} \lambda_n^{(k)} = \sum_{n=1}^{\infty} P\{T^{(k)}(\omega) = n\} = P(\bigcup_{n=1}^{\infty} \{T^{(k)}(\omega) = n\}) \leq 1.$$

Bewering: voor ieder nat. getal k geldt $\lambda_n^{(k+1)} = \sum_{i=1}^{n-1} \lambda_i^{(k)}.\lambda_{n-i}^{(1)}$.

Immers: $\lambda_n^{(k+1)} = P\{T^{(k+1)}(\omega) = n\}$

$$= P\{S_1(\omega)<k+1, S_2(\omega)<k+1,\ldots,S_{n-1}(\omega)<k+1, S_n(\omega) = k+1\}$$

$$= P(\bigcup_{i=1}^{n-1} \{S_1(\omega)<k,\ldots,S_{i-1}(\omega)<k, S_i(\omega)=k, S_{i+1}(\omega)<k+1,$$
$$\ldots,S_{n-1}(\omega)<k+1, S_n(\omega) = k+1\}$$

$$= \sum_{i=1}^{n-1} P\{S_1(\omega)<k,\ldots,S_{i-1}(\omega)<k, S_i(\omega)=k, S_{i+1}(\omega)<k+1,$$
$$\ldots,S_{n-1}(\omega)<k+1, S_n(\omega) = k+1\}$$

$$= \sum_{i=1}^{n-1} P\{S_1(\omega)<k,\ldots,S_{i-1}(\omega)<k, S_i(\omega)=k, \chi_{i+1}(\omega)<1,$$
$$\ldots,\chi_{i+1}(\omega)+\chi_{i+2}(\omega)+\ldots+\chi_{n-1}(\omega)<1, \chi_{i+1}(\omega)+\ldots+\chi_n(\omega) = 1\}$$

$$= \sum_{i=1}^{n-1} P(\underbrace{\{S_1(\omega)<k,\ldots,S_{i-1}(\omega)<k, S_i(\omega) = k\}}_{\text{gebeurtenis bepaald door de worpen 1 t/m i}} \cap$$

$$\underbrace{\{\chi_{i+1}(\omega)<1,\ldots,\chi_{i+1}(\omega)+\ldots+\chi_{n-1}(\omega)<1, \chi_{i+1}(\omega)+\ldots+\chi_n(\omega) = 1\}}_{\text{gebeurtenis bepaald door de worpen i+1 t/m n}})$$

$$= \sum_{i=1}^{n-1} P\{S_1(\omega)<k,\ldots,S_{i-1}(\omega)<k, S_i(\omega) = k\}.$$
$$P\{\chi_{i+1}(\omega)<1,\ldots,\chi_{i+1}(\omega)+\ldots+\chi_{n-1}(\omega)<1, \chi_{i+1}(\omega)+\ldots+\chi_n(\omega) = 1\}$$

$$= \sum_{i=1}^{n-1} \lambda_i^{(k)}.P\{S_1(\omega)<1, S_2(\omega)<1,\ldots,S_{n-i-1}(\omega)<1, S_{n-i}(\omega) = 1\}$$

$$= \sum_{i=1}^{n-1} \lambda_i^{(k)}.\lambda_{n-i}^{(1)}$$

Als we ter vereenvoudiging van de notatie $L^{(1)}(s)$ voorstellen door $L(s)$ dan geldt: voor ieder nat. getal k is $L^{(k+1)}(s) = L^{k+1}(s)$.
Immers: uit de laatste Bewering volgt $\{\lambda_n^{(k+1)}\} = \{\lambda_n^{(k)}\} * \{\lambda_n^{(1)}\}$ en uit het lemma van pag. 34 volgt dan $L^{(k+1)}(s) = L^{(k)}(s).L(s)$, zodat $L^{(k+1)}(s) = (L^{(k-1)}(s).L(s)).L(s)$ etc.

Bewering: $L(s) = ps + qsL^2(s)$.
Immers:

$$\lambda_1^{(1)} = P\{T^{(1)}(\omega) = 1\} = P\{S_1(\omega) = 1\} = P\{\chi_1(\omega) = 1\} = p$$

$$\begin{aligned}
\lambda_n^{(1)} &= P\{T^{(1)}(\omega) = n\} \\
&= P\{S_1(\omega)<1, S_2(\omega)<1, \ldots, S_{n-1}(\omega)<1, S_n(\omega) = 1\} \\
&= P\{\chi_1(\omega)=-1, \chi_2(\omega)<2, \chi_2(\omega)+\chi_3(\omega)<2, \ldots, \chi_2(\omega)+\ldots+\chi_{n-1}(\omega)<2, \\
&\quad \chi_2(\omega)+\ldots+\chi_n(\omega)=2\} \\
&= P\{\chi_1(\omega)=-1\}.P\{\chi_2(\omega)<2, \chi_2(\omega)+\chi_3(\omega)<2, \ldots, \chi_2(\omega)+\ldots+\chi_{n-1}(\omega)<2, \\
&\quad \chi_2(\omega)+\ldots+\chi_n(\omega)=2\} \\
&= P\{\chi_1(\omega)=-1\}.P\{S_1(\omega)<2, S_2(\omega)<2, \ldots, S_{n-2}(\omega)<2, S_{n-1}(\omega)=2\} \\
&= q\ \lambda_{n-1}^{(2)}.
\end{aligned}$$

$$\begin{aligned}
\Longrightarrow L(s) &= \lambda_1^{(1)} s + \sum_{n=2}^{\infty} \lambda_n^{(1)} s^n \\
&= ps + \sum_{n=2}^{\infty} q\lambda_{n-1}^{(2)} s^n \\
&= ps + qs \sum_{m=1}^{\infty} \lambda_m^{(2)} s^m \\
&= ps + qsL^{(2)}(s)\ .
\end{aligned}$$

Bewering: voor $s \neq 0$ en $-1 \leq s \leq +1$ geldt $L(s) = \frac{1}{2qs}\,(1-(1-4pqs^2)^{\frac{1}{2}})$.
Immers:
$L(s)$ is een machtreeks die convergeert voor $-1 \leq s \leq +1$; dat betekent o.a. dat voor iedere $-1 \leq s \leq +1$ $L(s)$ een ondubbelzinnig bepaald getal is. Bovendien geldt dat $L(s)$ op $-1 \leq s \leq +1$ een kontinue funktie is. Substitueren we $s=0$ in

de machtreeks dan vinden we $L(0) = 0$. Uit de laatste Bewering volgt dat voor iedere $s\neq 0$ $L(s)$ aan de vierkantsvergelijking $qsL^2(s) - L(s) + ps = 0$ voldoet, dus voor $s\neq 0$ is $L(s) = \frac{1}{2qs}\,(1\pm(1-4pqs^2)^{\frac{1}{2}})$.
Daar $\lim_{s\downarrow 0} \frac{1}{2qs}\,(1+(1-4pqs^2)^{\frac{1}{2}}) = \infty \neq L(0)$, moet dus

$$L(s) \text{ gelijk zijn aan } \frac{1}{2qs}\,(1-(1-4pqs^2)^{\frac{1}{2}}).$$

Bewering: $\lambda_{2n}^{(1)} = 0$ en $\lambda_{2n-1}^{(1)} = \frac{1}{2n-1}\binom{2n-1}{n} p^n q^{n-1}$.
Immers:

$\lambda_n^{(1)}$ is per definitie de coëfficient van s^n in de machtreeksontwikkeling van $L(s) = \frac{1}{2qs}\,(1-(1-4pqs^2)^{\frac{1}{2}})$

$$= \frac{1}{2qs}\,\Big(1-\sum_{n=0}^{\infty}\binom{\frac{1}{2}}{n}(-4pqs^2)^n\Big)$$

$$= -\frac{1}{2qs}\sum_{n=1}^{\infty}\binom{\frac{1}{2}}{n}(-4pq)^n s^{2n} \qquad \text{zodat}$$

$$\lambda_{2n}^{(1)} = 0 \text{ en } \lambda_{2n-1}^{(1)} = -\frac{1}{2q}\binom{\frac{1}{2}}{n}(-4pq)^n = \frac{1}{2n-1}\binom{2n-1}{n} p^n q^{n-1}.$$

We hebben hiermee berekend hoe groot de kans is dat Peter na n spelen voor het eerst op winst komt te staan. Het is niet moeilijk om in te zien dat de winst of het verlies van Peter na een even aantal worpen altijd even is en dus dat $\lambda_{2n}^{(1)}$ gelijk aan nul moet zijn.
Wat ons wellicht meer interesseert is te weten hoe groot de kans is dat Peter <u>ooit</u> op winst komt te staan.

Bewering: De kans dat Peter ooit op winst komt te staan is gelijk aan $\min(1,\frac{p}{q})$.

Immers: de gevraagde kans is gelijk aan:

$$P\Big(\bigcup_{n=1}^{\infty}\{T^{(1)}(\omega) = n\}\Big) = \sum_{n=1}^{\infty} P\{T^{(1)}(\omega) = n\}$$

$$= \sum_{n=1}^{\infty}\lambda_n^{(1)}$$

$$= L(1)$$

$$= \frac{1}{2q}\left(1-(1-4pq)^{\frac{1}{2}}\right)$$

$$= \frac{1}{2q}\left(1-|p-q|\right)$$

$$= \begin{cases} 1 \text{ als } p \geq q \\ \frac{p}{q} \text{ als } p<q \end{cases}$$

Bewering: de kans dat Peter ooit op een winst van k (k nat. getal) gulden komt te staan is gelijk aan $\min(1,(\frac{p}{q})^k)$

Immers:

de gevraagde kans is gelijk aan:

$$P\left(\bigcup_{n=1}^{\infty}\{T^{(k)}(\omega) = n\}\right) = \sum_{n=1}^{\infty} P\{T^{(k)}(\omega) = n\}$$

$$= \sum_{n=1}^{\infty} \lambda_n^{(k)}$$

$$= L^k(1)$$

$$= \min(1,(\tfrac{p}{q})^k)$$

We noemen het spel tussen Peter en Paul "eerlijk" en de Bernoulli wandeling "symmetrisch" als $p = q = \frac{1}{2}$.
In dit geval is de kans dat Peter ooit een winst van k gulden zal hebben gelijk aan 1. Dat we verwachten mogen dat het aantal worpen voorafgaande aan de worp waarna Peter voor het eerst 1 gulden gewonnen heeft groot zal zijn, blijkt als we $\mathcal{E}(T^{(1)})$ berekenen, voor $p = q = \frac{1}{2}$.

Lemma: Zij de machtreeks $A(s) = \sum_{n=0}^{\infty} a_n s^n$ convergent voor $-1<s<+1$.

Als $a_n \geq 0$ voor alle n dan geldt:
$\lim_{s\uparrow 1} A(s) = \sum_{n=0}^{\infty} a_n$, waarbij we afspreken dat we $\sum_{n=0}^{\infty} a_n = \infty$ stellen
als $\lim_{N\to\infty} \sum_{n=0}^{N} a_n = \infty$.

Bewijs: voor $\sum_{n=0}^{\infty} a_n$ geldt òf $\sum_{n=0}^{\infty} a_n < \infty$ òf $\sum_{n=0}^{\infty} a_n = \infty$.

Als $\sum_{n=0}^{\infty} a_n < \infty$ dan is de uitspraak een bijzonder geval van de continuïteitsstelling voor machtreeksen van Abel.

Voor $\sum_{n=0}^{\infty} a_n = \infty$: Kies een totaal willekeurig getal c; bij c is er een nat. getal N zodanig dat $\sum_{n=0}^{N} a_n > c$.

$\sum_{n=0}^{N} a_n s^n$ is een polynoom en dus een kontinue funktie van s

$\Longrightarrow$ er is een $\delta>0$ zodanig dat $\sum_{n=0}^{N} a_n s^n > c$ mits $1-\delta<s\leq 1$

$\Longrightarrow \sum_{n=0}^{\infty} a_n \geq \sum_{n=0}^{\infty} a_n s^n \geq \sum_{n=0}^{N} a_n s^n > c$ mits $1-\delta<s<1$.

Bewering: $\mathcal{E}(T^{(1)}) = \infty$ voor $p = q = \frac{1}{2}$.

Immers: $\mathcal{E}(T^{(1)}) = \sum_{n=1}^{\infty} n\lambda_n^{(1)}$.

Daar $\sum_{n=1}^{\infty} \lambda_n^{(1)} s^n$ convergeert voor $-1\leq s\leq +1$, convergeert $\sum_{n=1}^{\infty} n\lambda_n^{(1)} s^{n-1}$ voor $-1<s<+1$ en dus volgens bovenstaand lemma

$$\sum_{n=1}^{\infty} n\lambda_n^{(1)} = \lim_{s\uparrow 1} \sum_{n=1}^{\infty} n\lambda_n^{(1)} s^{n-1}$$

$$= \lim_{s\uparrow 1} L'(s) = \lim_{s\uparrow 1} \frac{d}{ds}\left(\frac{1}{s}(1-(1-s^2)^{\frac{1}{2}}\right) = \infty.$$

Opgaven:

1) Bereken $\mathcal{E}(T^{(k)})$ voor $p = q = \frac{1}{2}$;
2) Voor $p>q$ geldt $\sum_{n=1}^{\infty} \lambda_n^{(1)} = 1$; bereken voor dit geval $\mathcal{E}(T^{(1)})$.
3) Toon aan dat we $L^{(-k)}(s)$ (k nat. getal) verkrijgen door in $L^k(s)$ p en q te verwisselen.

Definitie: $U(\omega) = \min \{m|m\geq 1, S_m(\omega) = 0\}$

$U(\omega)$ noemen we de "terugkeertijd".

$\{U(\omega) = n\}$ is de verzameling van ω's waarvoor we in de n-de stap voor het eerst in het "uitgangspunt" terugkeren. Dit is een gebeurtenis die bepaald wordt door de eerste n worpen en is dus een meetbare verzameling. Zij $f_n = P\{U(\omega) = n\}$ en laat verder $F(s) = \sum_{n=1}^{\infty} f_n s^n$ de genererende funktie zijn van de rij $\{f_n\}_{n=1}^{\infty}$.

Bewering: de machtreeks $F(s)$ convergeert voor $-1 \leq s \leq +1$

Immers: $\{U(\omega) = n\} \cap \{U(\omega) = m\} = \emptyset$ voor $n \neq m$

$$\Longrightarrow \sum_{n=1}^{\infty} f_n = \sum_{n=1}^{\infty} P\{U(\omega) = n\} = P\left(\bigcup_{n=1}^{\infty} \{U(\omega) = n\}\right) \leq 1$$

Bewering: $f_n = q\lambda_{n-1}^{(1)} + p\lambda_{n-1}^{(-1)}$

Immers:

$$\begin{aligned}
f_n &= P\{U(\omega) = n\} \\
&= P\{S_n(\omega) = 0,\ S_m(\omega) \neq 0 \text{ voor } 1 \leq m < n\} \\
&= P(\{S_1(\omega)=-1, S_2(\omega)<0, \ldots, S_{n-1}(\omega)<0, S_n(\omega)=0\} \cup \{S_1(\omega)=+1, S_2(\omega)>0, \ldots, \\
&\qquad S_{n-1}(\omega)>0, S_n(\omega)=0\}) \\
&= P\{\chi_1(\omega)=-1, \chi_2(\omega)<1, \ldots, \chi_2(\omega)+\ldots+\chi_{n-1}(\omega)<1, \chi_2(\omega)+\ldots+\chi_n(\omega)=1\} \\
&\quad +P\{\chi_1(\omega)=+1, \chi_2(\omega)>-1, \ldots, \chi_2(\omega)+\ldots+\chi_{n-1}(\omega)>-1, \chi_2(\omega)+\ldots+\chi_n(\omega)=-1\} \\
&= P\{\chi_1(\omega)=-1\}.P\{\chi_2(\omega)<1, \ldots, \chi_2(\omega)+\ldots+\chi_{n-1}(\omega)<1, \chi_2(\omega)+\ldots+\chi_n(\omega)=1\} \\
&\quad +P\{\chi_1(\omega)=+1\}.P\{\chi_2(\omega)>-1, \ldots, \chi_2(\omega)+\ldots+\chi_{n-1}(\omega)>-1, \chi_2(\omega)+\ldots+\chi_n(\omega)=-1\} \\
&= q.\lambda_{n-1}^{(1)}+p\lambda_{n-1}^{(-1)}.
\end{aligned}$$

Bewering: $F(s) = 1-(1-4pqs^2)^{\frac{1}{2}}$

Immers:

$$\begin{aligned}
F(s) &= \sum_{n=1}^{\infty} f_n s^n = \sum_{n=2}^{\infty} f_n s^n \quad (\text{immers } f_1 = P\{U(\omega) = 1\} = P(\emptyset) = 0\ ) \\
&= \sum_{n=2}^{\infty} (q\lambda_{n-1}^{(1)} + p\lambda_{n-1}^{(-1)})s^n \\
&= qs \sum_{n=2}^{\infty} \lambda_{n-1}^{(1)}\, s^{n-1} + ps \sum_{n=2}^{\infty} \lambda_{n-1}^{(-1)}\, s^{n-1} \\
&= qs\, L(s) + ps\, L^{(-1)}(s) \\
&= \frac{qs}{2qs}\,(1-(1-4pqs^2)^{\frac{1}{2}}) + \frac{ps}{2ps}\,(1-(1-4pqs^2)^{\frac{1}{2}}) \\
&= 1-(1-4pqs^2)^{\frac{1}{2}}
\end{aligned}$$

Bewering: $f_{2n+1} = 0$ en $f_{2n} = \frac{2}{2n-1} \binom{2n-1}{n} p^n q^n$

Immers:

$$F(s) = 1-(1-4pqs^2)^{\frac{1}{2}}$$

$$= 1- \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-4pqs^2)^n$$

$$= - \sum_{n=1}^{\infty} \binom{\frac{1}{2}}{n} (-4pq)^n s^{2n}$$

$$\Longrightarrow f_{2n+1} = 0 \text{ en } f_{2n} = -\binom{\frac{1}{2}}{n} (-4pq)^n = \frac{2}{2n-1} \binom{2n-1}{n} p^n q^n .$$

Bewering: de kans dat Peter en Paul ooit quitte spelen is gelijk aan $\min(1,2p,2q)$.

Immers:

de gevraagde kans is gelijk aan

$$P\left(\bigcup_{n=1}^{\infty} \{U(\omega) = n\}\right) = \sum_{n=1}^{\infty} P\{U(\omega) = n\}$$

$$= \sum_{n=1}^{\infty} f_n$$

$$= F(1)$$

$$= 1-|p-q|$$

$$= \begin{cases} 1 & \text{als} \quad p = q = \frac{1}{2} \\ 2p & \text{als} \quad p<q \\ 2q & \text{als} \quad p>q \end{cases}$$

Uit de laatste Bewering volgt dat als een deeltje zich beweegt als in een symmetrische Bernoulli wandeling, het deeltje met kans 1 naar zijn uitgangspunt zal terugkeren.

Bewering:
voor $p = q = \frac{1}{2}$ geldt $\mathcal{E}(U) = \infty$

Immers:

$$\mathcal{E}(U) = \sum_{n=1}^{\infty} nf_n = \lim_{s\uparrow 1} F'(s) = \infty$$

Ruïneringskansen.

We hebben in het voorafgaande berekend dat als $p \geq q$, Peter met kans 1 ieder bedrag hoe groot ook kan winnen. Dat dit weleens lang kan duren hebben we gezien aan $\mathcal{E}(T^{(k)})$. Waar we niet opgelet hebben is het verlies dat Peter eventueel kan lijden alvorens hij een van te voren vastgesteld bedrag wint. Daarom gaan we nu veronderstellen dat Peter begint met een kapitaal ter grootte van z guldens en we vragen ons af hoe groot de kans is dat hij zijn beginkapitaal verliest voordat hij a-z guldens gewonnen heeft. Dit komt overeen met het berekenen van de kans dat Peter geruïneerd wordt als hij een beginkapitaal heeft van z guldens, zijn tegenspeler Paul een beginkapitaal heeft van a-z guldens en het spel pas dan gestaakt wordt als één der spelers geruïneerd is.

Stellen we de kans dat Peter geruïneerd wordt gelijk aan q_z *), dan geldt voor $1<z<a-1$ dat q_z voldoet aan de differentievergelijking

$$q_z = pq_{z+1} + qq_{z-1} \tag{1}$$

Immers:

Peter zal na de eerste worp z+1 of z-1 guldens te verliezen hebben al naar gelang hij wint of verliest in de eerste worp.

Voor z=1 geldt: $q_1 = pq_2 + q$. Om deze vergelijking aan te passen aan (1) voeren we in het getal

$$q_0 = 1 \tag{2}$$

Dan geldt $q_1 = pq_2 + qq_0$.

Voor $z = a-1$ geldt: $q_{a-1} = qq_{a-2}$;

introduceer $\quad q_a = 0 \qquad (3)$

dan geldt $\quad q_{a-1} = pq_a + qq_{a-2}$.

---

*) Peter wordt precies na n worpen geruïneerd voor de ω's uit $\{-z<S_1(\omega)<a-z,\ldots,-z<S_{n-1}(\omega)<a-z, S_n(\omega) = -z\}$; deze verzameling is meetbaar, noem hem A_n. Dan is ook $\bigcup_{n=1}^{\infty} A_n$ meetbaar en $q_z = P(\bigcup_{n=1}^{\infty} A_n)$

Voor $z = 1,2,\ldots,a-1$ voldoen de kansen q_z aan de differentievergelijking (1). We zullen ze gaan berekenen door (1) op te lossen. Zoals gebruikelijk zullen we allerlei oplossingen van (1) aanduiden met q_z.
In de theorie van de differentievergelijkingen worden (2) en (3) randvoorwaarden genoemd.

Bewering: Als q_z en q_z^* oplossingen zijn van (1) én

$$\begin{cases} q_0 = q_0^* \\ q_1 = q_1^* \end{cases}$$

dan zijn de oplossingen q_z en q_z^* identiek.

Immers:

$$\begin{aligned} q_1 &= pq_2 + qq_0 \\ \Longrightarrow q_2 &= \frac{1}{p}(q_1 - qq_0) \\ &= \frac{1}{p}(q_1^* - qq_0^*) = q_2^* \\ \Longrightarrow q_3 &= \frac{1}{p}(q_2 - qq_1) \\ &= \frac{1}{p}(q_2^* - qq_1^*) = q_3^* \quad \text{etc.} \end{aligned}$$

Bewering: voor $p \neq q$ is iedere oplossing van (1) van de vorm

$$C_1 + C_2\left(\frac{q}{p}\right)^z \qquad (C_1 \text{ en } C_2 \text{ zijn constanten}).$$

Immers:
Door substitutie in (1) vinden we dat $q_z = 1$ en $q_z = (\frac{q}{p})^z$ oplossingen van (1) zijn. Ook iedere lineaire combinatie van deze oplossingen is een oplossing van (1). Dus voor ieder paar constanten (C_1,C_2) is $C_1 + C_2(\frac{q}{p})^z$ een oplossing.
We moeten nu nog aantonen dat als we een totaal willekeurige oplossing van (1) hebben, noem hem maar q_z^*, dat er dan constanten C_1 en C_2 bestaan zodanig dat $q_z^* = C_1 + C_2(\frac{q}{p})^z$.

Maar als de constanten C_1 en C_2 voldoen aan:

$$\begin{aligned} C_1 + C_2 &= q_0^{*} \\ C_1 + C_2 . \frac{q}{p} &= q_1^{*} \end{aligned} \qquad (4)$$

dan volgt uit de laatste Bewering dat $q_z^{*} = C_1 + C_2(\frac{q}{p})^z$.
Het niet strijdige stelsel ($p \neq q$) van lineaire vergelijkingen in de onbekenden C_1 en C_2 is altijd op te lossen.

Bewering: Er is precies één oplossing van de differentievergelijking (1) die voldoet aan de randvoorwaarden (2) en (3) en wel

$$q_z = \frac{(\frac{q}{p})^a - (\frac{q}{p})^z}{(\frac{q}{p})^a - 1} . \qquad (5)$$

Immers:
er is precies één stel constanten C_1, C_2 dat voldoet aan

$$C_1 + C_2 = 1 = q_0$$

$$C_1 + C_2(\frac{q}{p})^a = 0 = q_a$$

en $\quad C_1 + C_2(\frac{q}{p})^z$ is gelijk aan (5) (ga dit na).

Opgaven:

1) Leid af dat voor $p = q = \frac{1}{2}$ de oplossing van (1) die voldoet aan (2) en (3) gelijk is aan: $q_z = 1 - \frac{z}{a}$
2) Bereken de kans dat Peter één gulden wint voordat hij zijn beginkapitaal ter grootte van 999 gulden verliest in een eerlijk spel ($p = q = \frac{1}{2}$).

10. De binomiale verdeling.

Laten we de spelregels van het spel tussen Peter en Paul van de vorige paragraaf eens wat veranderen: Bij munt krijgt Peter nog steeds een gulden, doch bij kruis behoeft hij niet meer te betalen.

Overeenkomstig wijzigen we het bijbehorende WH-veld:

$$\Omega_1 = \{\omega | \mathbb{N} \xrightarrow{\omega} \{1,0\}\}$$

De definities van $\mathfrak{A}_1$ en P_1 verkrijgen we uit die van $\mathfrak{A}_0$ en P_0 door overal -1 te vervangen door 0.

De definitie van S_n blijft ongewijzigd; deze functie stelt nog steeds voor:

De totale winst van Peter na precies n worpen.

Dit is nu echter gelijk aan het aantal keren munt in n worpen.

S_n is een stochastische functie op het WH-veld $<\Omega_1, \mathfrak{A}_1, P_1>$. We vragen nu naar de bijbehorende verdelingsfunctie (deze hangt ook af van n en p).

<u>Definitie</u>: $b(k;n,p) = P_1\{S_n = k\}$

d.w.z. $b(k;n,p)$ is de kans op precies k keren munt (succes) in n worpen. In vraagstuk 5, pag. 36 is reeds gevraagd $b(k;n,p)$ te berekenen.

Het resultaat was: $b(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k}$.

Door deze uitkomst begrijpen we waarom we hier spreken van een binomiale verdeling.

Bij betrekkelijk kleine waarden van n is een directe berekening van $b(k;n,p)$ niet moeilijk.

Zowel uit practische als uit theoretische overwegingen stellen we echter vooral belang in het gedrag van deze functie voor grote n.
In het volgende zullen we twee belangrijke benaderingsmethoden behandelen die voor dit doel zijn ontwikkeld.

Om niet telkens aan WH-velden $<\Omega, \mathfrak{A}, P>$ te hoeven refereren voeren we een stelsel van discrete WH-velden (§1) in die als referentiekader van de volgende beschouwingen kunnen dienen.

Zij $0<p<1$ en n een natuurlijk getal;
dan is $\mathcal{B}(n,p) = <\mathbb{N}, f>$ met $f(k) = b(k;n,p)$
(als $k>n$ dan is $b(k;n,p) = 0$).

Beschouwen we op $\mathcal{B}(n,p)$ nu de stochastische functie X met $X(i) = i$ dan hebben we

$$P\{X = k\} = \sum_{X(i)=k} f(i) = f(k) = b(k;n,p),$$

zodat de verdelingsfunctie van X op $\mathcal{B}(n,p)$ precies de binomiale verdeling is die behoort bij n en p.

We zullen in het volgende ook de bij X behorende genormaliseerde functie X^* nodig hebben.

In §6 is gedefinieerd $X^* = \dfrac{X - \mu}{\sigma}$, met

$$\mu = \mathcal{E}(X) = \sum_{k=1}^{n} k\binom{n}{k} p^k (1-p)^{n-k} = np$$

(beschouw $x.\frac{\partial}{\partial x}(x+y)^n$)
en

$$\sigma = (\mathcal{E}(X^2) - n^2p^2)^{\frac{1}{2}} = (p^2n^2 - p^2n + pn - p^2n^2)^{\frac{1}{2}} =$$

$$(p(1-p)n)^{\frac{1}{2}} = \sqrt{npq} \qquad (q = (1-p))$$

dus $X^*(i) = \dfrac{i - np}{\sqrt{npq}}$

11. De poissonverdeling.

Vele problemen in de WH-theorie leiden tot de vraag naar een schatting van de frequenties van zeer onwaarschijnlijke gebeurtenissen bij zeer lange series proefnemingen.

In de terminologie van de vorige paragraaf wil dit zeggen:

p klein en n groot.

<u>Voorbeeld</u>. Bepaalde natuurprocessen

(b.v. het uiteenvallen van radioactieve materie)

vinden zo plaats dat de kans dat in een tijdvak een verandering (b.v. uiteenvallen van een kern) optreedt alleen afhangt van de lengte van dat tijdvak, en wel in eerste benadering lineair.

Zij $p(t)$ nu de kans dat in een tijdvak van lengte t zo'n verandering plaatsvindt, dan is voor een constante λ: $p(t) = \lambda t$

i.h.b. is dan $p_n \overset{def}{=} p(\frac{1}{n}) = \frac{\lambda}{n}$.

De kans dat in een tijdvak van lengte 1 k veranderingen optreden noemen we $p(k,\lambda)$.

Nemen we nu n zo groot dat praktisch in geen tijdvak ter lengte $\frac{1}{n}$ nog meer dan één verandering optreedt, dan geldt dat $p(k,\lambda)$ bij benadering gelijk is aan $b(k;n,\frac{\lambda}{n})$. (Beschouw een tijdvak als een proefneming en het optreden van een verandering als succes.)
Als de veranderingen van zodanige aard zijn dat er geen twee tegelijk op kunnen treden hebben we dus:

$$\lim_{n\to\infty} b(k;n,\frac{\lambda}{n}) = p(k,\lambda)$$

Dat deze limiet bestaat is uit de voorgaande beschouwing wel intuitief duidelijk. We kunnen dit ook door berekening aantonen, waarbij we voor $p(k,\lambda)$ een eenvoudige uitdrukking vinden:

$$b(k;n,\frac{\lambda}{n}) = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{n!}{k!(n-k)!} \cdot \frac{\lambda^k}{(n-\lambda)^k} \left(1 - \frac{\lambda}{n}\right)^n =$$

$$= \frac{(n-k+1)(n-k+2)\ldots(n-1)n}{(n-\lambda).(n-\lambda)\ldots(n-\lambda)} \cdot \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^k$$

Waaruit volgt

$$\lim_{n\to\infty} b(k;n,\frac{\lambda}{n}) = \frac{\lambda^k}{k!} e^{-\lambda}$$

De functie $p(k,\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}$ noemen we de poisson distributie functie.

We willen nu nog meten wat de orde van grootte is van de fout die we maken door $b(k;n,\frac{\lambda}{n})$ te benaderen met $p(k,\lambda)$.

Een elementaire berekening leert ons dat de fout die we maken door

$$\frac{(n-k+1)(n-k+2)\ldots(n-1)n}{(n-\lambda)(n-\lambda)\ldots(n-\lambda)(n-\lambda)}$$

door 1 te vervangen van de orde $\frac{1}{n}$ is.

Verder geldt:

$$\log\left(1 - \frac{\lambda}{n}\right)^n = n \log\left(1 - \frac{\lambda}{n}\right) = -\lambda - \frac{\lambda^2}{2n} - \ldots$$

dus $\left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda - \frac{\lambda^2}{2n} - \ldots}$ wat via de reeksontwikkeling voor e^x oplevert:

$$\left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda} + O\left(\frac{1}{n}\right)$$

We hebben dus

$$b(k;n,\frac{\lambda}{n}) = p(k,\lambda) + O\left(\frac{1}{n}\right) .$$

<u>Vraagstuk</u> Hoe groot is (bij benadering) de kans dat van 1000 mensen er precies 10 op 29 februari jarig zijn.

12. De formule van Stirling

In deze paragraaf geven wij een afleiding van de bekende formule van Stirling

$$n! \sim n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi} \qquad \text{voor } n\to\infty$$

en een ordeschatting van de fout die gemaakt wordt bij vervanging van het linkerlid door het rechter.

Hulpstelling 1

$$\int_0^{\frac{\pi}{2}} \sin^n x \, dx = \frac{n-1}{n} \int_0^{\frac{\pi}{2}} \sin^{n-2} x \, dx \qquad (n \geq 2)$$

Bewijs: We noteren $I_n = \int_0^{\frac{\pi}{2}} \sin^n x \, dx$;

$$\text{dan is } I_n = -\sin^{n-1} x \cos x \Big|_0^{\frac{\pi}{2}} + (n-1) \int_0^{\frac{\pi}{2}} \cos^2 x \sin^{n-2} x \, dx$$

$$= (n-1) \int_0^{\frac{\pi}{2}} (1 - \sin^2 x) \sin^{n-2} x \, dx$$

$$= (n-1)\, I_{n-2} - (n-1)\, I_n$$

dus $nI_n = (n-1)\, I_{n-2}$ q.e.d.

We merken op: $I_0 = \frac{\pi}{2}$, $I_1 = 1$

Hieruit volgt

$$I_{2n} = \frac{(2n-1)(2n-3)\ldots 1}{2n(2n-2)\ldots 2} \cdot \frac{\pi}{2} = \frac{(2n)!}{(2^n n!)^2} \cdot \frac{\pi}{2}$$

$$I_{2n+1} = \frac{2n(2n-2)(2n-4)\dots 2}{(2n+1)(2n-1)\dots 3} = \frac{(2^n n!)^2}{(2n+1)(2n)!}$$

$$I_{2n-1} = \frac{(2n-2)\dots}{(2n-1)\dots} = \frac{(2^n n!)^2}{2n(2n)!}$$

<u>Hulpstelling 2</u> (Product van Wallis)

$$\lim_{n\to\infty} \frac{(2^n n!)^2}{(2n)!\sqrt{n}} = \sqrt{\pi}$$

<u>Bewijs</u>: Aangezien op $(0, \frac{\pi}{2})$ $0 < \sin x < 1$

is $\sin^{2n+1} x < \sin^{2n} x < \sin^{2n-1} x$,

dus ook $I_{2n+1} < I_{2n} < I_{2n-1}$

oftewel $\frac{(2^n n!)^2}{(2n+1)(2n)!} < \frac{(2n)!}{(2^n n!)^2} \cdot \frac{\pi}{2} < \frac{(2^n n!)^2}{2n(2n)!}$

Omgewerkt: $\frac{n}{n+\frac{1}{2}} < \left(\frac{(2n)!}{(2^n n!)^2}\right)^2 n\pi < 1$

Hieruit volgt onze bewering onmiddellijk.

<u>Opmerking</u>: $\int_0^x \log t \, dt = x(\log x - 1)$

<u>Bewijs</u>: Partiële integratie

<u>Stelling</u> (formule van Stirling)

$$n! = n^{n+\frac{1}{2}} e^{-n+\frac{1}{12n}+O(\frac{1}{n^2})} \cdot \sqrt{2\pi}$$

<u>Bewijs</u>: We proberen eerst $\log(n!) = \sum_{k=1}^{n} \log k$ te benaderen

Welke fout maken we als we deze vorm vervangen door de integraal

$$\int_0^n \log x \, dx = n(\log n - 1) \quad ?$$

We stellen $s_n = \log(n!) - n(\log n - 1)$ en beschouwen het gedrag van s_n voor $n \to \infty$. Daartoe berekenen we $s_n - s_{n-1}$.

$$s_n - s_{n-1} = \log n - n \log n + (n-1) \log(n-1) + 1$$

$$= 1 + (n-1) \log \left(1 - \frac{1}{n}\right) =$$

$$= 1 - (n-1) \left(\frac{1}{n} + \frac{1}{2n^2} + \frac{1}{3n^3} + \dots\right) =$$

$$= \frac{1}{2n} + \frac{1}{6n^2} + 0\left(\frac{1}{n^3}\right)$$

d.w.z. $s_n = \sum_{k=1}^{n} \left(\frac{1}{2k} + 0\left(\frac{1}{k^2}\right)\right)$

We weten dat $\lim_{n\to\infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{n} - \log n\right) = \gamma$, (de constante van Euler-Mascheroni) $\sum_{k=1}^{n} \frac{1}{2k}$ zal dus naar we mogen verwachten voor $n \to \infty$ hoogstens een constante van $\frac{1}{2} \log n$ verschillen.
Laten we daarom nu eens beschouwen:

$$S_n = s_n - \frac{1}{2} \log n = \log(n!) - n(\log n - 1) - \frac{1}{2} \log n$$

We berekenen weer

$$S_n - S_{n-1} = s_n - s_{n-1} + \frac{1}{2} \log\left(1 - \frac{1}{n}\right) =$$

$$= \frac{1}{2n} + \frac{1}{6n^2} + \frac{1}{2} \log\left(1 - \frac{1}{n}\right) + 0\left(\frac{1}{n^2}\right) ;$$

met behulp van de reeksontwikkeling van $\log\left(1 - \frac{1}{n}\right)$ vinden we

$$S_n - S_{n-1} = -\frac{1}{12n^2} + 0\left(\frac{1}{n^3}\right) .$$

Dus de rij S_n heeft een limiet S en $S - S_n = -\frac{1}{12} \sum_{k=n+1}^{\infty} \frac{1}{k^2} + 0\left(\frac{1}{n^2}\right)$

Uit de ongelijkheid $\frac{1}{k} - \frac{1}{k+1} = \frac{1}{k(k+1)} < \frac{1}{k^2} < \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}$ leiden we gemakkelijk af dat

$$\sum_{k=n+1}^{\infty} \frac{1}{k^2} = \frac{1}{n} + 0(\frac{1}{n^2})$$

dus $\qquad S - S_n = - \frac{1}{12n} + 0(\frac{1}{n^2})$

We hebben nu gevonden

$$\log(n!) = S + n(\log n - 1) + \frac{1}{2}\log n + \frac{1}{12n} + 0(\frac{1}{n^2})$$

Anders geformuleerd:

$$n! = e^S . n^{n+\frac{1}{2}} . e^{-n + \frac{1}{12n} + 0(\frac{1}{n^2})}$$

Rest ons nog e^S te berekenen.

Substitutie in de formule van Wallis levert

$$\sqrt{\pi} = \lim_{n\to\infty} \frac{2^{2n} e^{2s} n^{2n+1} e^{-2n+\frac{1}{6n}}}{e^S (2n)^{2n+\frac{1}{2}} e^{-2n+\frac{1}{24n}} . n^{\frac{1}{2}}} = \lim_{n\to\infty} \frac{e^S}{\sqrt{2}}$$

Hieruit volgt $e^S = \sqrt{2\pi}$ zodat onze stelling bewezen is.

<u>Gevolg 1</u>: $n! = n^n e^{-n}\sqrt{2\pi n} \cdot (1 + 0(\frac{1}{n}))$

Immers $e^x = 1 + x + 0(x^2)$ voor $x \to 0$.

<u>Gevolg 2</u>: $\binom{n}{k} = (\frac{n}{k})^k (\frac{n}{n-k})^{n-k} \sqrt{\frac{n}{2\ k(n-k)}}\ (1 + 0(\frac{1}{n}))$

Dit volgt direct door toepassing van gevolg 1.

13. De normale approximatie.

We zullen nu een benadering beschouwen van de verdeling van de §10 behandelde genormaliseerde stochastische functie X^*.
We vragen naar de kans dat X^* tussen twee gegeven waarden α en β ligt.

$$P\{\alpha < X^* < \beta\} = P\{np + \alpha\sqrt{npq} < X < np + \beta\sqrt{npq}\}.$$

Zij nu $U_{n,p} = (np + \alpha\sqrt{npq},\ np + \beta\sqrt{npq})$
dan is

$$P\{\alpha < X^* < \beta\} = \sum_{k \in U_{n,p}} P\{X=k\} = \sum_{k \in U_{n,p}} b(k;n,p)$$

We zoeken nu eerst een benadering van $b(k;n,p)$ onder de voorwaarde $k \in U_{n,p}$.
voor $n \to \infty$. (dit impliceert ook $k \to \infty$ en $n-k \to \infty$)
Hiertoe gebruiken we gevolg 2 uit §12.
Dit levert op

$$b(k;n,p) = \frac{(np)^k (nq)^{n-k}}{k^k (n-k)^{n-k}} \sqrt{\frac{n}{2\pi k(n-k)}}\ (1 + 0(\tfrac{1}{n}))$$

Stellen we nu $\delta_k = k - np$ dan is $k = np + \delta_k$ en $n-k = nq - \delta_k$.
Dus:

$$(1)\qquad b(k;n,p) = \frac{1}{(1 + \frac{\delta_k}{np})^{np+\delta_k}} \cdot \frac{1}{(1 - \frac{\delta_k}{nq})^{nq-\delta_k}} \sqrt{\frac{n}{2\pi(np+\delta_k)(nq-\delta_k)}} \cdot (1 + 0(\tfrac{1}{n}))$$

We zullen eerst aantonen dat

$$(2)\qquad \frac{n}{2\pi(np+\delta_k)(nq-\delta_k)} = \frac{1}{\sqrt{2\pi npq}}\ (1 + 0(\tfrac{1}{\sqrt{n}}))$$

Immers $\alpha\sqrt{npq} < \delta_k < \beta\sqrt{npq}$ zodat

$$(3)\qquad \delta_k = 0(\sqrt{n})$$

waaruit (2) direct volgt.

Verder geldt:

$$\log \frac{1}{(1+\frac{\delta_k}{np})^{np+\delta_k}} \quad \frac{1}{(1-\frac{\delta_k}{nq})^{nq-\delta_k}} =$$

$$(np+\delta_k)\log(1+\frac{\delta_k}{np}) + (nq-\delta_k)\log(1-\frac{\delta_k}{nq}) =$$

$$(np+\delta_k)\left(\frac{\delta_k}{np} - \frac{\delta_k^2}{2n^2p^2} + \frac{\delta_k^3}{3n^3p^3} - \ldots\right) -$$

$$(nq-\delta_k)\left(\frac{\delta_k}{nq} + \frac{\delta_k^2}{2n^2q^2} + \frac{\delta_k^3}{3n^3p^3} + \ldots\right).$$

Deze ontwikkeling is geldig voor voldoend grote n want $\frac{\delta_k}{np} = 0(n^{-1/2})$ evenzo $\frac{\delta_k}{nq} = 0(n^{-1/2})$.

Nemen we nu termen samen waarin δ_k tot dezelfde macht voorkomt dan krijgen we

$$\frac{\delta_k^2}{2n}\left(\frac{1}{p}+\frac{1}{q}\right) - \frac{\delta_k^3}{6n^2}\left(\frac{1}{p^2}-\frac{1}{q^2}\right) + \ldots ;$$

echter $\frac{\delta_k^3}{n^2} = 0(n^{-1/2})$ zodat

(4) $\quad b(k;n,p) = \frac{e^{-\frac{1}{2}x_k^2}}{\sqrt{2\pi npq}}\left(1+0(\frac{1}{\sqrt{n}})\right)$ voor $n\to\infty$ en $k \in U_{n,p}$ met $x_k = \frac{\delta_k}{\sqrt{npq}}$.

Dus

$$\sum_{k\in U_{n,p}} b(k;n,p) = \frac{1}{\sqrt{2\pi npq}} \sum_{k\in U_{n,p}} e^{-\frac{1}{2}x_k^2}\left(1+0(n^{-1/2})\right) \quad \text{voor } n\to\infty$$

Het aantal elementen in $U_{n,p}$ is $K = [(\beta-\alpha)\sqrt{npq}]$.

Zij nu voor $0 \leq j \leq K$, $y_j = \alpha + \frac{j}{\sqrt{npq}}$, $y_{K+1} = \beta$ en $k_j = [\alpha\sqrt{npq}]^+ + j$ dan is

$$\frac{1}{\sqrt{2\pi npq}} \sum_{k\in U_{n,p}} e^{-\frac{1}{2}x_k^2} = \frac{1}{\sqrt{2\pi}} \left(\sum_{j=0}^{K+1} e^{-\frac{1}{2}x_{k_j}^2}(y_{j+1}-y_j) - e^{-\frac{1}{2}x_{k_K}^2}(y_k-\beta)\right)$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{j=0}^{K+1} e^{-\frac{1}{2}x_{k_j}^2}(y_{j+1}-y_j) + 0(\frac{1}{n})$$

met $y_j \leq x_{k_j} \leq y_{j+1}$. Dit als Riemann-som beschouwend vinden we

$$\lim_{n\to\infty} P\{\alpha < X^* < \beta\} = \frac{1}{\sqrt{2\pi}} \int_\alpha^\beta e^{-\frac{t^2}{2}} dt.$$

De verdeling van de genormaliseerde functie X^* nadert dus voor $n\to\infty$ tot de "normale verdelingsfunctie":

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt .$$

Opmerking 1

De hier bewezen stelling wordt de limietstelling van De Moivre en Laplace genoemd. We zien dat de limietfunctie niet van p afhangt: De invloed van p (b.v. asymmetrie) op de genormaliseerde binomiale verdeling (d.i. de verdeling van X^*) is bij grote n blijkbaar verwaarloosbaar.

Opmerking 2

(5) is ook te schrijven als

(6) $$\lim_{n\to\infty} P\{\alpha < X^* < \beta\} = \Phi(\beta) - \Phi(\alpha)$$

Opmerking 3

De gegeven benaderingen gelden niet voor Bernouilli proefnemingen waarbij n afhangt van het behaalde resultaat, dus niet voor een spel waarmee een speler kan ophouden b.v. als hij genoeg gewonnen heeft of teveel verloren.

Vraagstuk: Stel dat er dagelijks tussen 8 en 9 uur 50.000 forensen van het Gooi naar Amsterdam reizen en gemiddeld 30% de trein neemt. (De kans dat een reiziger de trein neemt stellen we dus op 0,3) De spoorwegen willen het risico dat iemand pas na negenen een trein kan krijgen tot 1% beperken. Hoeveel treinstellen voor maximaal 260 passagiers moeten daartoe in dit tijdvak rijden? (als er m treinstellen rijden moet de kans dat er meer dan 250m passagiers komen 0,01 zijn).

<u>Stelling</u> $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = 1$

<u>Bewijs</u> $\left(\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx\right)^2 = \iint_{\mathbb{R}^2} e^{-\frac{x^2+y^2}{2}} dx\, dy =$

$$= \int_0^{2\pi} d\phi \int_0^{\infty} e^{-\frac{r^2}{2}} r dr = \int_0^{2\pi} d\phi = 2\pi$$

Tabel van de Normale Verdeling.

x	$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$	$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$
0,0	0,398.942	0,500.000
0,1	,396.952	,539.828
0,2	,391.043	,579.260
0,3	,381.388	,617.911
0,4	,368.270	,655.422
0,5	,352.065	,691.462
0,6	,333.225	,725.747
0,7	,312.254	,758.036
0,8	,289.692	,788.145
0,9	,266.085	,815.940

1,0	,241.971	,841.345
1,1	,217.852	,864.334
1,2	,194.186	,884.930
1,3	,171.369	,903.200
1,4	,149.727	,919.243
1,5	,129.518	,933.193
1,6	,110.921	,945.201
1,7	,094.049	,955.435
1,8	,078.950	,964.070
1,9	,065.616	,971.283
2,0	,053.991	,977.250
2,1	,043.984	,982.136
2,2	,035.475	,986.097
2,3	,028.327	,989.276
2,4	,022.395	,991.802
2,5	,017.528	,993.790
2,6	,013.583	,995.339
2,7	,010.421	,996.533
2,8	,007.915	,997.445
2,9	,005.953	,998.134
3,0	,004.432	,998.650
3,1	,003.267	,999.032
3,2	,002.384	,999.313
3,3	,001.723	,999.517
3,4	,001.232	,999.663
3,5	,000.873	,999.767
3,6	,000.612	,999.841
3,7	,000.425	,999.892
3,8	,000.292	,999.928
3,9	,000.199	,999.952
4,0	,000.134	,999.968
4,1	,000.089	,999.979
4,2	,000.059	,999.987
4,3	,000.039	,999.991
4,4	,000.025	,999.995
4,5	,000.016	,999.997

Opmerking: $\Phi(-x) = 1 - \Phi(x)$.