Hoofdstuk 4

Een inleiding in kansrekenen

1. Waarom kansrekenen voor criminologen?

Een criminoloog moet regelmatig beslissingen nemen onder onzekere omstandigheden. Bij de beslissing om een pedofiel vrij te laten is er sprake van een zekere kans dat de dader hervalt. Om deze kans op zijn waarde te schatten, moeten we weten wat een kans is. Een criminoloog die een uitspraak wilt doen over de onveiligheidsbeleving van de Vlaming, kan moeilijk elke Vlaming bevragen. Zes miljoen Vlamingen een vragenlijst doen invullen, is net iets van het goede te veel. Criminologen zullen daarom een steekproef nemen. Wetenschappelijke steekproeven moeten representatief zijn voor de bevolking. Steekproeven worden op verschillende manieren getrokken. Het belangrijkste is dat er een toevalsmechanisme achter de steekproef schuilt. Dit betekent dat elk element een zelfde kans heeft om getrokken te worden. Ook hier spelen kansen een rol en is het essentieel dat je een absoluut basisvocabularium en basisprincipes die schuilgaan achter de kansrekening begrijpt. Een criminoloog kan in de praktijk geconfronteerd worden met de vraag van een opdrachtgever om een steekproef te nemen van het cliënteel waar hij of zij mee werkt, met de vraag een bepaalde dienst te evalueren. Welke soort steekproef dient de criminoloog dan te nemen? We zullen verderop zien dat er nogal wat mogelijkheden bestaan. Toevalssteekproeven verschillen van elkaar in de kans op een bepaalde uitkomst. Je zou bijvoorbeeld elk element in de populatie een nummer kunnen geven en dan via toevalsmechanisme A (zonder teruglegging) je steekproef bepalen, maar je zou ook elke respondent die je getrokken hebt opnieuw de kans kunnen geven om getrokken te worden. Waarom bestaan er zulke verschillen en waarom zouden we kiezen voor op maat gemaakte steekproeven? In dit hoofdstuk wordt hierop een antwoord geboden. Aan de hand van **drie kansdefinities** maken we duidelijk hoe een elementaire kans kan bepaald worden. Daarna geven we enkele kansregels om kansen te berekenen in iets ingewikkeldere situaties. Dit wordt gedaan via de introductie van het concept voorwaardelijke kans, i.e. de kans op een gebeurtenis onder een bepaalde andere conditie.

2. Kansdefinities

Experimenten, waarvan wordt geëist dat ze een onbeperkt aantal keren kunnen worden herhaald onder gelijkblijvende omstandigheden, worden kansexperimenten genoemd. Een steekproef trekken is ook zo een situatie. Een **kans** is steeds een getal tussen 0 en 1 (een proportie), maar vaak drukken we deze uit in een percentage. De notatie van een kans is de letter **P**. De kans om een zes te gooien met één dobbelsteen wordt genoteerd als **P(6)** en wordt uitgesproken als de kans op een 6. De vraag die we ons stellen, is: kunnen we bepaalde kansen berekenen? En zo ja, hoe doen we dat dan? Welke is de kans om te winnen in een loterij? Welke is de kans om in de gevangenis te belanden? Welke is de kans dat men slachtoffer wordt van een geweldsdelict als men in een grote stad woont?

In een aantal gevallen is dit heel eenvoudig. Je berekent de **kansdefinitie van Laplace**. De kansdefinitie van Laplace is het aantal uitkomsten waarin je geïnteresseerd bent, gedeeld door het aantal mogelijke uitkomsten van een kansexperiment. De kansdefinitie van Laplace is een theoretische of objectieve kansdefinitie.

De formule die daar bij hoort is: **N(A)/N** waarbij

A= uitkomst van het kansexperiment (de gebeurtenis) waar je in geïnteresseerd bent

N(A)= het aantal uitkomsten waar je in geïnteresseerd bent

N= het aantal uitkomsten

Als je met een dobbelsteen gooit en je wilt weten hoe groot de kans is dat je een even getal gooit als je een keer mag gooien, ga je als volgt te werk:

A zijn de uitkomsten waarin je geïnteresseerd bent: 2,4,6

Het aantal uitkomsten N(A) waarin je geïnteresseerd bent is 3 (je bent in die drie gebeurtenissen geïnteresseerd). Het aantal mogelijke uitkomsten is echter 1, 2, 3, 4, 5 en 6. Er zijn dus zes mogelijkheden terwijl er slechts drie interessant zijn. De kans om even te gooien is: P(A)= N (A)/ N ofwel 3/6= 50%.

Het is niet altijd mogelijk om de objectieve of theoretische kans te bepalen. Als dat onmogelijk is, kan men wel de relatieve frequentie bepalen. Deze noemen we ook vaak de experimentele kans en is gebaseerd op eigen onderzoek.
De relatieve frequentie is te bepalen door het kansexperiment te herhalen en vervolgens te kijken hoe vaak de uitkomst die je interesseert zich voordoet ten opzichte van het totaal aantal herhaalde kansexperimenten.

Hierbij hoort de formule:

P(A)= n(A)/n

Waarbij:

n(A)= het aantal uitkomsten waarin je geïnteresseerd bent bij herhaling van het experiment

n= het aantal herhalingen van het experiment

Hier komen we al vrij dicht in de buurt van wat beleidsmakers soms van criminologen willen weten. In veel “real life” situaties is een kans niet op logischerwijze te bepalen. Als je zelf een nieuwe weekendbijverdienste in een kroeg zoekt en je wilt weten hoe groot de kans is dat er ‘s weekends in de kroeg waar je overweegt te gaan werken, vechtpartijen uitbreken, dan kan je daarover cijfers verzamelen. Je kan natuurlijk niet elk weekend cijfers verzamelen. Als je maar enkele weekends observaties uitvoert, dan zal je eigen onderzoek weinig precieze uitkomsten genereren. Maar als je er een jaar lang elk weekend observaties maakt, dan weet je toch veel preciezer hoe groot de kans is dat de gebeurtenis waarin je geïnteresseerd bent, effectief kan plaatsvinden. Als je dertig weekends hebt geobserveerd en we hebben op 3 weekends problemen geobserveerd, dan is de relatieve kans (volgens jouw bevindingen) 3/30 of 1/10 of 10%. Kansen worden hier gedefinieerd in termen van veelvuldig herhalen. Hier speelt wat men noemt **de experimentele wet**: naarmate het aantal herhalingen van een toevalsproces toeneemt, zullen de kansen van de elementen van S (de steekproef) zich meer en meer stabiliseren. Bij een groot aantal herhalingen wordt een stabiele waarde beschouwd als de kans dat element uit S zich voordoet. Criminologen zien steekproefgegevens als uitkomst van een kansproces!

Het is niet altijd mogelijk om een kansexperiment uit te voeren. Wanneer je mensen op straat hoort praten over hun eigen impressies, over kansen, dan spreken we over **subjectieve kansen**, in tegenstelling tot objectieve kansen. De subjectieve kans is de eigen inschatting en is dus gebaseerd op de perceptie van het individu. De kans dat student Jan Janssens later een succesvol gevangenisdirecteur zal worden, staat niet geschreven op het diploma van de student, maar kan de persoon zelf trachten in te schatten. Het resultaat van deze eigen inschatting is de **subjectieve kans**.

3. Kansregels

Om kansen te berekenen, bestaan er verschillende regels. In deze syllabus dienen volgende kansregels gekend te zijn: **de algemene somregel**, **de speciale somregel**, **de algemene productregel**, **de speciale productregel**, **de complementregel.** Deze regels zijn van belang omdat kansrekenen de wiskundige beschrijving van toevalsfenomenen is. Inzicht in de absolute basis hiervan is noodzakelijk om te begrijpen hoe we aan inductieve statistiek doen, of anders gezegd hoe we als criminologen in staat zijn om op basis van één steekproef uitspraken te doen over een onderzoekspopulatie. We gaan in deze syllabus enkel en slechts enkel de basisregels toelichten.

De algemene somregel luidt als volgt:

P(A of B)= P(A) + P(B) – P(A en B)

We zijn hier geïnteresseerd in de twee gebeurtenissen A en B en we willen weten of één van deze twee gebeurtenissen afzonderlijk plaatsvindt, <u>maar niet gezamenlijk</u>.

Voorbeeld uit een kaartspel (52 kaarten):
13 harten
13 ruiten
13 schoppen
13 klaveren

In iedere set van kaarten zit een vrouw. Er zijn dus 4 vrouwen. Je wilt weten hoe groot de kans is dat je of een vrouw of een hartenkaart trekt.

P(A)= Probabiliteit harten= 13/52

P(B)= Probabiliteit vrouw= 4/52

P (harten en vrouw)= 1/52

Merk op dat de hartenvrouw twee maal geteld wordt, daarom moet deze er één maal afgehaald worden. A en B hebben namelijk één zaak gemeenschappelijk.
De kans op harten of vrouw is dan: 13/52 + 4/52 – 1/52= 16/52 of 30.77%

De speciale somregel luidt als volgt:

P(A of B)= P(A) + P(B) waarbij geldt dat A en B niets met elkaar gemeen hebben.

Opnieuw een voorbeeld uit een kaartspel: als je wilt weten wat de kans is op harten of ruiten, dan tel je de beide kansen op.

P(harten)= 13/52
P(ruiten)= 13/52
P(harten of ruiten)= 13/52 + 13/52 = 50%.
Precies omdat er geen dubbeltellingen zijn, kunnen we deze beide kansen optellen.

De algemene productregel
In de realiteit is er vaak overlap tussen twee gebeurtenissen. Daarom moet het begrip **voorwaardelijke kans** geïntroduceerd worden. De voorwaardelijke kans op een bepaalde uitkomst is de kans op een bepaalde uitkomst als je al gedeeltelijke informatie hebt over de uitkomst van een kansexperiment. Stel dat je met een dobbelsteen gooit en je wil weten hoe groot de kans is dat je hoogstens drie gooit. Omdat je weet dat er evenveel getallen kleiner of gelijk aan drie zijn dan groter dan drie, heb je voorkennis en kan je de voorwaardelijke kans berekenen. **De voorwaardelijke kans is de kans op gebeurtenis A, onder conditie van gebeurtenis B.** In de bivariate statistiek zullen we vaak met conditionele kansen geconfronteerd worden, bijvoorbeeld de kans dat iemand een veelpleger is, gegeven het feit dat men jongen is, versus de kans dat iemand veelpleger is, gegeven de kans dat iemand meisje is.

De notatie van een **voorwaardelijke of conditionele kans** luidt:

$$\mathbf{P(B \mid A)= P\ (A\ en\ B)\ /\ P\ (A)}$$

Dit is de kans op B gegeven A. Als we deze regel toepassen en we willen weten wat de kans is op het gooien van een getal dat niet hoger is dan drie (de kans op B), gegeven het feit dat het een even getal moet zijn (de kans op A) dan bekomen we:

P(A en B)= de probabiliteit dat een getal even is EN hoogstens drie is= 1/6. Enkel het getal twee is immers kleiner of gelijk aan drie en even.

P(A)= de probabiliteit dat een getal even is= 3/6

De voorwaardelijke kans= 1/6 gedeeld door 3/6 of 1/6 * 6/3= 6/18 of 1/3 of 33.33%.

De speciale productregel

In veel gevallen is vooraf ontvangen informatie die je krijgt bij een kansexperiment niet relevant voor het kansprobleem. Zo is de opmerking dat het in New York file rijden is op zaterdag weinig relevant om de kans te kennen dat je op zaterdag in Brussel file zal rijden. Het effect is dan dat P (B│A) = P (B). De kans dat A en B samen voorkomen is dan **P (B) * P (A)**. Dit noemen we de **speciale productregel**.

We passen deze regel toe aan de hand van een voorbeeld. Stel dat we een vaas hebben met zes zwarte en vier rode ballen. We halen er toevalsgewijs een bal uit en leggen die daarna terug. Methodologen noemen dit een steekproef met teruglegging. De kans dat we de eerste keer een zwarte bal trekken, heeft in zulke situatie geen enkele invloed op de kans dat we de tweede keer een zwarte bal trekken. We geven nog een voorbeeld: we trekken een steekproef met teruglegging van eerstejaarsstudenten criminologie. We weten dat er 60% meisjes en 40% jongens zijn. Bij de eerste trekking wordt een meisje getrokken. Er is teruglegging dus hetzelfde meisje kan opnieuw geselecteerd worden. Wat is de kans dat opnieuw een meisje getrokken wordt? Door de teruglegging is de kans opnieuw 60%. De kans dat de tweede keer opnieuw een meisje geselecteerd wordt is 60%*60%= 36%.

De complementregel

Het kan voorkomen dat een kans lastig te berekenen is vanwege de complexiteit van de vraagstelling. In een aantal gevallen kan het daarom handig zijn om naar het tegenovergestelde probleem te kijken. Als dit probleem eenvoudiger op te lossen is, kan men aan de hand van de complementregel de kans berekenen.

P(A) = 1-P (complement van A)

Als de kans om een zes te gooien met één dobbelsteen 1/6 is, dan is de kans om geen zes te gooien 1- (1/6)= 5/6. Als je weet dat de kans op het geslaagd zijn in de eerste zittijd van het eerste bachelorjaar criminologie 1/10 is, dan weet je tegelijk dat je 90% kans hebt om niet geslaagd te zijn in de eerste zittijd van het eerste bachelorjaar criminologie.

4. Permutaties en combinaties

Bij kansrekenen moet er vaak veel geteld worden om bij het kansexperiment het aantal elementen te bepalen waar je in geïnteresseerd bent. Een hulp hierbij is het aantal permutaties en combinaties. Het aantal **permutaties** is **het aantal manieren waarop je een aantal verschillende objecten ten opzichte van elkaar** kan plaatsen. De volgorde is dus van belang. Als we onszelf de vraag stellen op hoeveel manieren we de letters A, B, C ten opzichte van elkaar kunnen plaatsen, dan zien we:

ABC, ACB, BAC, BCA, CAB, CBA

Dit zijn zes mogelijkheden. De redenering bij permutaties: voor de eerste letter zijn er drie mogelijkheden, namelijk A, B of C. Als we voor de eerste letter een keuze gemaakt hebben, dan blijven er voor de tweede letter nog 2 mogelijkheden over. De derde letter ligt dan vast, er blijft immers maar één letter meer over. **Het aantal permutaties van n objecten wordt weergegeven als n! (spreek uit; n faculteit) en is: n*(n-1)…*1.**

vb: 5!=5*4*3*2*1=120

Bij combinaties zijn er bij een groep objecten twee subgroepen. Binnen iedere subgroep zijn de elementen niet van elkaar te onderscheiden. We tellen het aantal manieren om de objecten ten opzichte van elkaar neer te zetten.

We tellen het aantal combinaties dat er mogelijk is met de letters A, A, B, B en B.

AABBB, ABABB, ABBAB, ABBBA, BAABB, BABAB, BABBA, BBAAB, BBABA, BBBAA

We zien dat er tien combinaties zijn. Dit hadden we ook als volgt kunnen vinden: als alle letters verschillend zijn, zijn er 5!= 120 mogelijkheden. Er zijn 2 A's. Het verwisselen van twee A's levert een dubbeltelling op. De 2 A's zijn op 2!= 2 manieren te verwisselen. We moeten delen door deze dubbeltelling. Er zijn 3 B's. Het verwisselen van die B's levert ook dubbeltellingen op. In totaliteit kunnen we bij het verwisselen van de 3 B's 3!= 6 dubbeltellingen vinden. We moeten ook door deze dubbeltellingen delen.

De berekening is: 5!/(2!*3!)= 120/12=10

Het aantal combinaties van n (in het voorbeeld 5) elementen met een subgroep van k (2) gelijke elementen en een tweede subgroep van n-k (5-2=3) gelijke elementen is gelijk aan 10.

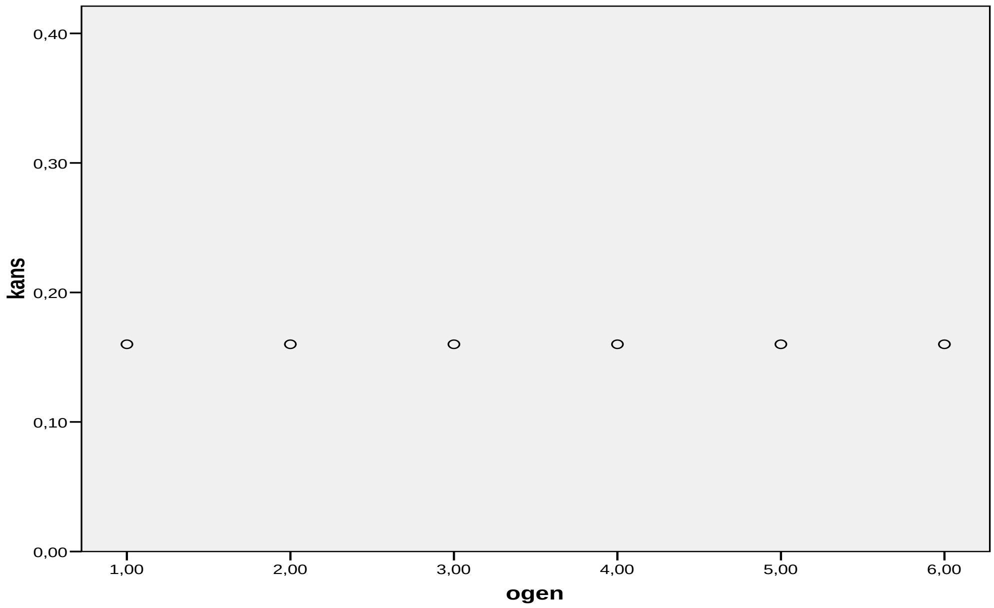
Het aantal combinaties van *k* elementen uit een verzameling van *n* elementen wordt formeel genoteerd als de binomiaalcoëfficiënt $\binom{n}{k}$ (*spreek uit*: *n* over *k*). Dit wordt ook als volgt genoteerd: $\binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}$ Deze binomiaalcoëfficiënt is nodig voor het berekenen van kansen dat willekeurige gebeurtenissen voorkomen in bepaalde combinaties.

5. Kansvariabelen en de binomiale verdeling

Een kansvariabele geeft aan in welk getal een kansexperiment resulteert. De notatie van een kansvariabele is bijvoorbeeld $\underline{\mathbf{k}}$ of $\underline{\mathbf{x}}$. Het streepje onder de letters verwijst naar het feit dat we te maken hebben met een kansexperiment en dat niet bij voorbaat vast staat wat de uitkomst zal zijn. Zo kunnen we kijken naar $\underline{k}$: het aantal ogen van een dobbelsteen bij eenmaal gooien met de uitkomsten (1, 2, 3, 4, 5, 6). In criminologisch onderzoek kijken we bijvoorbeeld naar de kans op recidive nadat 30 personen een alternatieve straf hebben gevolgd. Kansen kunnen berekend worden voor **continue** en **discrete** variabelen. Het voorbeeld van de ogen van de dobbelsteen en het voorbeeld van de kans op recidive zijn voorbeelden van een discrete kansverdeling. Drie personen van de dertig kunnen recidiveren, maar dat kunnen er ook vier zijn. Tussenliggende waarden zijn onmogelijk. Als we de kansen willen berekenen van de lengte van de files in Brussel, dan kan die file 3 kilometer zijn, maar ook 3.21 of 4.26. De kansen op iedere uitkomst in een kansexperiment is weer te geven met een kansfunctie. Kijken we naar het voorbeeld van k: het aantal ogen van een dobbelsteen bij eenmaal gooien, dan zijn de kansen op de uitkomsten:

$P(\underline{k}=1)= 1/6$; $P(\underline{k}=2)= 1/6$; $P(\underline{k}=3)= 1/6$; $P(\underline{k}=4)= 1/6$; $P(\underline{k}=5)=1/6$; $P(\underline{k}=6)=1/6$

Figuur: kansverdeling voor het gooien van ogen met één dobbelsteen



We kunnen de uitkomsten van een kansfunctie cumuleren:

P(k≤1)= 1/6
P(k≤2)= 2/6
P(k≤3)= 3/6
P(k≤4)= 4/6
P(k≤5)= 5/6
P(k≤6)= 6/6 of 1

We noemen dit overzicht de **verdelingsfunctie**. Een verdelingsfunctie is **het overzicht van de uitkomsten van een kansvariabele met de hierbij behorende gecumuleerde kansen**. Een kenmerk van een kansvariabele is de verwachte waarde. De verwachte waarde is de som van de uitkomsten vermenigvuldigd met de kans op iedere uitkomst.

In een formule wordt de verwachte waarde als volgt weergegeven:

$E\,\underline{k} = \sum P(\underline{k}=k)*k$

Bij het voorbeeld met de dobbelsteen is dit:

$E\,\underline{k} = 1/6*1 + 1/6*2 + 1/6*3 + 1/6*4 + 1/6*5 + 1/6*6 = 3,5$.

Dit betekent dat als je heel vaak (en in theorie oneindig vaak) met een dobbelsteen gooit en het gemiddelde berekent van je resultaten, je uiteindelijk 3,5 zal vinden als gemiddelde.

Tevens kan de standaardafwijking bij een kansvariabele bepaald worden. De standaardafwijking bij een discrete kansvariabele kan berekend worden met de formule:

$$\sigma = \sqrt{\sum P(\underline{k} = k)(k - E\underline{k})^2}$$

In het voorbeeld van de dobbelsteen is dit 1,71. Deze waarde vind je ook door heel vaak (oneindig vaak) het kansexperiment te herhalen en vervolgens de standaardafwijking ervan te berekenen.

6. De binomiale verdeling

In kwantitatief criminologisch onderzoek zijn kansvariabelen van groot belang. We geven opnieuw het voorbeeld van recidive-onderzoek. We willen weten hoe groot de kans is dat niemand tot iedereen die een alternatieve sanctie heeft gekregen, hervalt in oude gewoonten of recidiveert. M.a.w. we willen eigenlijk weten hoe effectief een behandelingsprogramma is. De kansvariabele die voor ons van belang is in het recidivevoorbeeld, is k: het aantal personen uit een groep van 5 personen die een alternatieve maatregel hebben gekregen, die na de alternatieve maatregel toch hervalt.
De mogelijke uitkomsten zijn:

$P(\underline{k}=0)$; $P(\underline{k}=1)$; $P(\underline{k}=2)$; … ; $P(\underline{k}=5)$

Hoe worden deze kansen nu berekend? Hoe groot is $P(\underline{k} = 0)$? Met andere woorden, hoe groot is de kans dat geen van de personen hervalt? Stel dat we uit een grootschalig onderzoek weten dat 40% van de personen die de alternatieve maatregel kregen opgelegd, hervalt. Hoe zit het dan bij onze groep van 5 personen? De kans dat iemand niet hervalt is dus 60%. Alle vijf de personen mogen in ons voorbeeld niet hervallen. We kunnen dus de speciale productregel toepassen om de kans te berekenen. Het feit dat één persoon hervalt heeft statistisch niets te maken met het feit dat een andere persoon hervalt.

De kans in onze groep is dus: $P(k= 0)= 0.40° * 0.60^5= 0.07776$

Wat als we willen weten wat de kans is dat precies één persoon hervalt?: $P(\underline{k} = 1)$. Deze persoon benoemen we met de letter J, alle andere met N. Dit geeft ons verschillende combinatiemogelijkheden:

JNNNN (De eerste persoon hervalt, de overigen niet)

NJNNN (Enkel de tweede persoon hervalt)

NNJNN (Enkel de derde persoon hervalt)

NNNJN (Enkel de vierde persoon hervalt)

NNNNJ (Enkel de laatste persoon hervalt)

Of anders: de formule van de binomiaalcoefficiënt $\binom{n}{k}$ = n!/(k!(n-k!) toegepast op ons voorbeeld:

5!/(1!*4!)= 5

Er zijn dus 5 manieren waarop we precies één persoon kunnen vinden die hervalt.

Hoe groot is dan P(JNNNN)? M.a.w. hoe groot is de kans dat de combinatie JNNNN voorvalt? Deze kans vinden we door de toepassing van de speciale productregel: $0.40^1 * 0.60^4 = 0.0518$. Het voorkomen van één van de andere vier combinaties kent dezelfde kans.

Wanneer we de kans P(<u>k</u>=1) willen kennen. Anders gezegd, wanneer de combinatiemogelijkheid niet van belang is, maar we willen weten hoe groot de kans is dat een willekeurig iemand recidiveert, moeten we de speciale productregel toepassen <u>maal het aantal manieren waarop we de k</u> (in ons voorbeeld: het aantal personen uit een groep van 5 personen die een alternatieve maatregel hebben gekregen, die na de alternatieve maatregel toch hervalt) <u>kunnen verdelen over de n waarnemingen</u>. M.a.w. we gebruiken de binomiaalcoëfficiënt om dit aantal te vinden.

De kans dat een willekeurig iemand recidiveert is dan:

P (<u>k</u> = 1) : $(0.40^1 * 0.60^4) * 5 = 0.2592$

Op dezelfde manier kunnen we de kans berekenen dat exact de volgende combinatie voorkomt: JJNNN (= De eerste twee personen recidiveren, de drie laatste niet).

De kans P (JJNNN) $=0.40^2 * 0.60^3 = 0.0346$.

Het aantal combinaties van 2 uit 5 (‘vijf over twee’) is:

5!/(2!*3!) = 10 mogelijkheden

De kans dat twee willekeurige personen recidiveren is dan:

P (k = 2) : ($0.40^2 * 0.60^3$) * 10 = 0.3456

En 3…

P (k = 3) : ($0.40^3 * 0.60^2$) * 10 = 0.2304

P (k = 4) : ($0.40^4 * 0.60^1$) * 5 = 0.0768

P (k = 5) : ($0.40^5 * 0.60^0$) = 0.01024

Als we de kansen cumuleren, krijgen we het volgende:

P (k ≤ 0) = P (k = 0) = 0.07776

P (k ≤ 1) = P (k ≤ 0) + (k = 1) = 0.07776 + 0.2592 = 0.33696

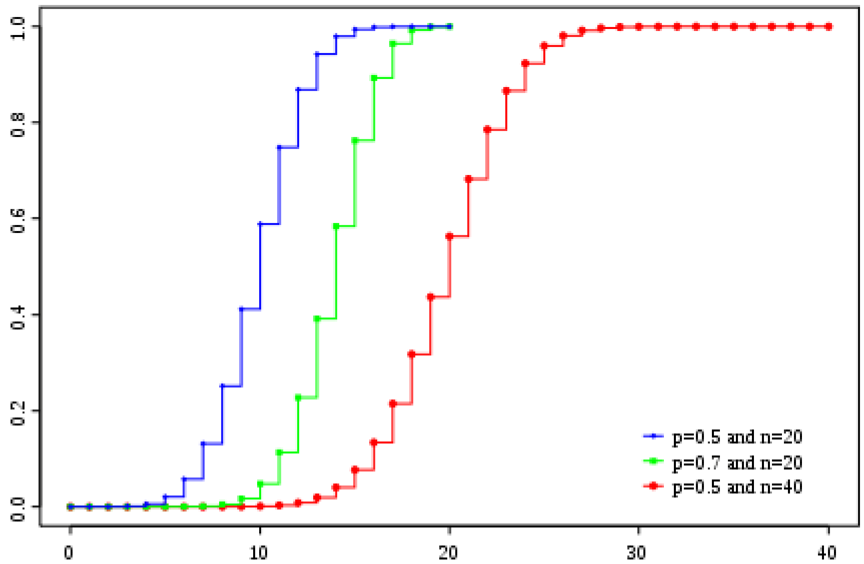
P (k ≤ 2) = P (k ≤ 1) + (k = 2) = 0.33696 + 0.3456 = 0.68256

P (k ≤ 3) = P (k ≤ 2) + (k = 3) = 0.68256 + 0.2304 = 0.91296

P (k ≤ 4) = P (k ≤ 3) + (k = 4) = 0.91296 + 0.0768 = 0.98976

P (k ≤ 5) = P (k ≤ 4) + (k = 5) = 0.98976 + 0.01024 = 1

Figuur: Cumulatieve binomiale verdelingen



7. De binomiale verdeling gaat over in een normale verdeling

De binomiale verdeling gaat over in een normale verdeling. We geven hieronder voorbeelden. Hoe meer tentamens en hoe meer studieresultaten, hoe meer combinaties van voldoenden en onvoldoenden er mogelijk zijn. Stel dat er slechts twee mogelijke uitslagen bij een tentamen zijn: Voldoende (V) en onvoldoende (O). Stel dat de kansen als volgt verdeeld zijn: de kans op een voldoende is 60% en op een onvoldoende 40%, of: P(V) = 0,6 en P(O) = 0,4
Bij één tentamen zijn de volgende uitkomsten mogelijk:

P(V) = 0,6
P(O) = 0,4

Bij twee tentamens zijn de volgende uitkomsten mogelijk:
P(VV) = P(V) * P(V) = 0.6 * 0.6 = 0.36
P(OV) = P(O) * P(V) = 0.4 * 0.6 = 0.24
P(VO) = P(V) * P(O) = 0.6 * 0.4 = 0.24
P(OO) = P(O) * P(O) = 0.4 * 0.4 = 0.16

Kans op voldoende bij twee tentamens (twee 'trekkingen'; n = 2):
Kans op 0 voldoenden = P(OO) = 0.16
Kans op 1 voldoende = P(OV) + P(VO) = 0.24 + 0.24 = 0.48 of anders: $(0.40^1 * 0.60^1)*2$ = 0.48
Kans op 2 voldoenden = P(VV) = 0.60^2 = 0.36

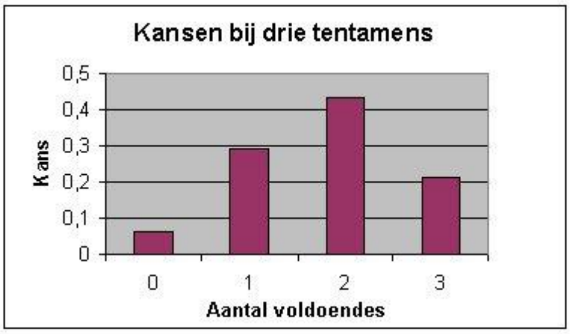
Kansen bij drie tentamens:
P(OOO) = 0,4 * 0,4 * 0,4 = 0,064
P(OOV) = 0,4 * 0,4 * 0,6 = 0,096
P(OVO) = 0,4 * 0,6 * 0,4 = 0,096
P(VOO) = 0,6 * 0,4 * 0,4 = 0,096
P(OVV) = 0,4 * 0,6 * 0,6 = 0,144
P(VOV) = 0,6 * 0,4 * 0,6 = 0,144
P(VVO) = 0,6 * 0,6 * 0,4 = 0,144
P(VVV) = 0,6 * 0,6 * 0,6 = 0,216

Kans op 0 voldoenden = P(OOO) = 0.064

Kans op 1 voldoende = P(OOV) + P(OVO) + P(VOO) = 3 * 0,096 = 0,288

Kans op 2 voldoenden = P(OVV) + P(VOV) + P(VVO) = 3 * 0,144 = 0,432

Kans op 3 voldoenden = P(VVV) = 0,216



Kansen bij vier tentamens (n = 4):

P(VVVV) = 0,6 * 0,6 * 0,6 * 0,6 = 0,1296

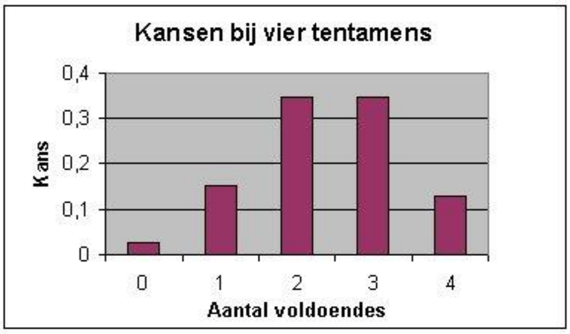
P(VVVO) = 0,6 * 0,6 * 0,6 * 0,4 = 0,0864

P(VVOV) = 0,6 * 0,6 * 0,4 * 0,6 = 0,0864

P(VOVV) = 0,6 * 0,4 * 0,6 * 0,6 = 0,0864

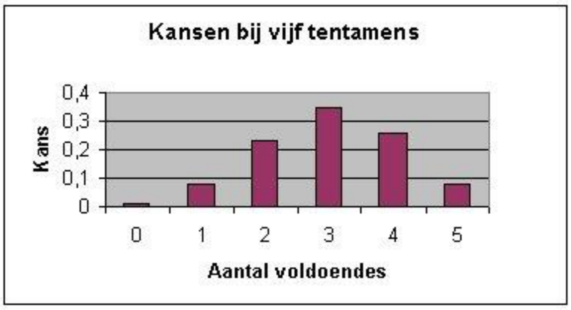
etc. tot: P(OOOO) = 0,4 * 0,4 * 0,4 * 0,4 = 0,0256

Aantal voldoendes	Kans
0	0,0256
1	0,1536
2	0,3456
3	0,3456
4	0,1296



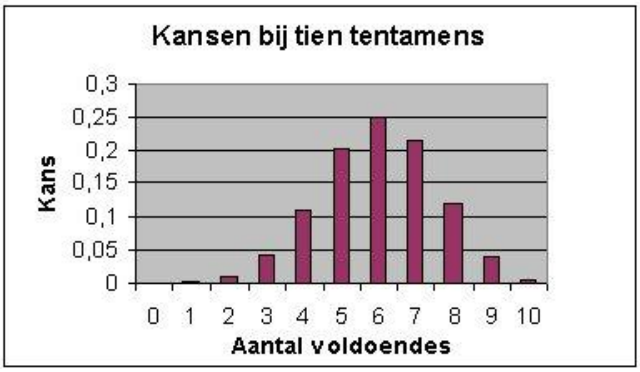
Kansen bij vijf tentamens (n = 5):

Aantal voldoendes	Kans
0	0,0102
1	0,0768
2	0,2304
3	0,3456
4	0,2592
5	0,0778

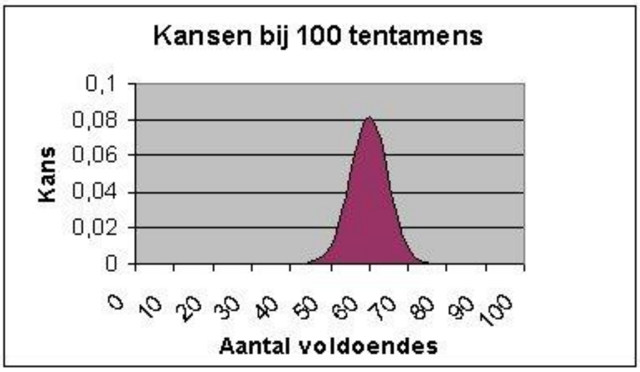


Kansen bij tien tentamens (n = 10):

Aantal voldoendes	Kans
0	0,000100768
1	0,001603456
2	0,010600778
3	0,0425
4	0,1115
5	0,2007
6	0,2508
7	0,2150
8	0,1209
9	0,0403
10	0,0060



Kansen bij 100 tentamens (n = 100):



Wat blijkt uit deze oefening? Naarmate het aantal tentamens toeneemt, benadert de kansverdeling de normale verdeling meer en meer. De grafische afbeelding van de normale verdeling is de Gauss-curve die we in een vorig hoofdstuk besproken hebben.

8. Waarom is de binomiale verdeling zo belangrijk in kwantitatief criminologisch onderzoek?

Veel kenmerken waarin de criminoloog geïnteresseerd is, volgen een binomiale verdeling. Het al of niet plaatsvinden van criminele gebeurtenissen zoals slachtofferschap, is zo een voorbeeld. Deze kenmerken zijn categorische variabelen, dichotomieën, bvb 1 = slachtoffer, 0 = geen slachtoffer. Wanneer we aan statistische inferentie doen, met name het veralgemenen van steekproefresultaten naar een bredere populatie, zullen we gebruik maken van de kenmerken van deze verdeling om uitspraken te doen over de veralgemeenbaarheid van onze bevindingen. Dit komt aan bod in de volgende delen.

9. Leerdoelen

Dit deel heeft tot doel de studenten een aantal elementaire kansregels aan te leren. Het begrip kansdefinitie dient goed gekend te zijn. De verschillende regels, met name de algemene somregel, de speciale somregel, de algemene productregel, de speciale productregel en de complementregel dienen gekend te zijn. Het is van belang te weten in welke situatie welke kansregel van toepassing is. Dit is belangrijk met betrekking tot het begrijpen van steekproefuitkomsten. Vervolgens werd in dit hoofdstuk aandacht besteed aan permutaties en combinaties. Deze beide begrippen dienen goed van elkaar te worden onderscheiden. Permutaties en combinaties dienen zelf te kunnen worden uitgerekend en evenals de kans op een bepaalde uitkomst.