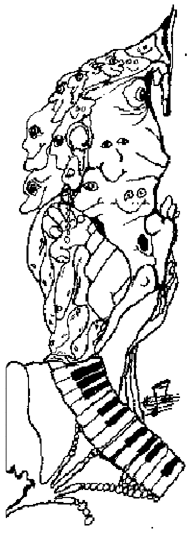


Willem van Ravenstein

Lesbrief Poisson-verdeling

"Life is good for only two things, discovering mathematics and teaching mathematics."
Simeon Poisson

Inhoudsopgave

Voorkennis

"De Poissonverdeling is genoemd naar Siméon Poisson die deze kansverdeling ontdekte en samen met zijn statistische theorie in 1838 publiceerde in zijn werk'Recherches sur la probabilité des jugements en matières criminelles et matière civile'."
http://nl.wikipedia.org/wiki/Poissonverdeling

De binomiale verdeling

We beschouwen n onafhankelijke experimenten met elk experiment een kans van p op 'succes'. De stochast X, die het totaal aantal 'successen' voorstelt, heeft een **binomiale verdeling** met parameters p en n.

Er geldt: $\mathrm{P}(\mathrm{X}=\mathrm{k})=\binom{\mathrm{n}}{\mathrm{k}}\cdot\mathrm{p}^{\mathrm{k}}\cdot(1-\mathrm{p})^{\mathrm{n}-\mathrm{k}}$

Voor de binomiale verdeling geldt: E(X)=n·p en Var(X)=n·p(1-p)

Telproblemen

Veel kansproblemen hebben te maken met **tellen**. Stel jezelf, voor je begint, de volgende **twee vragen**:

1. Is het met of zonder teruglegging?
2. Is de volgorde belangrijk?

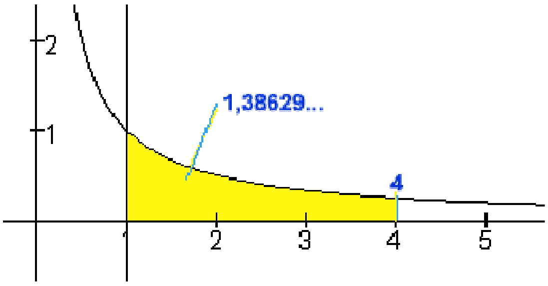
Dit levert vervolgens 4 verschillende soorten telproblemen op:

		Met terugleggen?	
		Nee	**Ja**
Volgorde belangrijk?	**Ja**	**Permutaties** $(n)_k=\dfrac{n!}{(n-k)!}$ **faculteitsboom**	**Rangschikkingen met herhaling** $aantal=n^k$ **machtsboom**
	Nee	**Combinaties** $\binom{n}{k}=\dfrac{n!}{(n-k)!\cdot k!}$ **ja-nee rooster**	**Herhalingscombinaties** $aantal=\binom{n-1+k}{k}$

Definities

We definiëren de functie **ln(x)** door: $\ln(x) = \int_1^x \frac{1}{t}dt$ voor $x \in \mathbb{R}$ met $x > 0$

De definitie van ln(x) heeft dus veel te maken met de oppervlakte onder de grafiek van $f(x) = \frac{1}{x}$ vanaf 1. Hieronder zie je de oppervlakte voor x=4. Dit is dan ongeveer 1,38629...



De functie **ln(x)** is continu en strikt stijgend op $< 0, \rightarrow >$. De **inverse** functie heet de **exponentiële functie.** Deze functie is gedefinieerd op $\mathbb{R}$ met als bereik $< 0, \rightarrow >$.

Notatie: $y = e^x$

Het getal 'e' definiëren we als het unieke getal waarvoor $\int_1^e \frac{1}{t}dt = 1$

$e \approx 2{,}718281828...$

Handig om te weten:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + ... \qquad \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e \qquad \lim_{n\to\infty}\left(1+\frac{b}{n}\right)^n = e^b$$

Hoofdstuk 1 - wiskundige afleiding van de Poissonverdeling[1]

De Poisson-verdeling kan je opvatten als een ‘limiet-geval’ van de binomiale verdeling. Als n groot is en n·p neemt een vaste waarde aan dan kan je de kans op een bepaalde gebeurtenis berekenen met de formule van de Poisson-verdeling: $P(X=k)=e^{-\lambda}\cdot\dfrac{\lambda^k}{k!}$

Deze verdeling wordt alleen bepaald door de verwachtingswaarde λ. De standaardafwijking is gelijk aan de wortel uit de verwachtingswaarde ($\sigma=\sqrt{\lambda}$).

In dit hoofdstuk leiden we de formule voor de Poissonverdeling af.

De binomiale verdeling: $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$ voor $k=0,1,\ldots,n$

We kiezen $\lambda=n\cdot p$ zodat $p=\dfrac{\lambda}{n}$. Invullen in de formule voor de binomiale verdeling geeft:

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

, zodat:

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

We gaan nu kijken wat er gebeurd als n groot wordt. We beginnen deze keer maar ’s achteraan:

I.

De factor $\dfrac{1}{\left(1-\frac{\lambda}{n}\right)^k}$ gaat bij vaste k naar 1 als $n\to\infty$

II.

Als $n\to\infty$ dan gaat bij een vaste waarde van k de factor $\dfrac{n!}{n^k(n-k)!}$ ook naar 1.

Immers: $\dfrac{n!}{n^k(n-k)!}=\dfrac{n(n-1)(n-2)\ldots(n-k+1)}{n^k}=\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\left(1-\frac{3}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)$

III.

Wat te doen met de factor $\left(1-\frac{\lambda}{n}\right)^n$? We weten al dat $\lim\limits_{n\to\infty}\left(1+\frac{b}{n}\right)^n=e^b$ dus de factor $\left(1-\frac{\lambda}{n}\right)^n$ gaat naar $e^{-\lambda}$ als $n\to\infty$.

IV.

We komen uiteindelijk uit op: $P(X=k)=e^{-\lambda}\cdot\dfrac{\lambda^k}{k!}$

[1] Dit hoofdstuk kan bij eerste lezing worden overgeslagen.☺

Hoofdstuk 2 - voorbeelden

Voorbeeld 1

In een bepaald gebied zijn er gemiddeld 4 blikseminslagen per jaar.
Bereken de kans op 0, 1, 2, 3, 4, 5, 6 en meer dan 6 blikseminslagen per jaar.

Uitwerking

We stellen vast $\lambda = 4$. Met $P(X = k) = e^{-4} \cdot \dfrac{4^k}{k!}$ kan je dan de kansen uitrekenen:

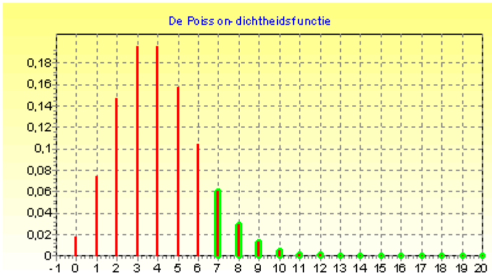
$$P(X = 0) = e^{-4} \cdot \frac{4^0}{0!} \approx 0{,}0183$$

$$P(X = 1) = e^{-4} \cdot \frac{4^1}{1!} \approx 0{,}0733$$

$$P(X = 2) = e^{-4} \cdot \frac{4^2}{2!} \approx 0{,}1465$$

...

$$P(X > 6) \approx 0{,}1107$$




```
poissonpdf(4,1)
      .0732625556
poissonpdf(4,2)
      .1465251111
1-poissoncdf(4,6
)
      .1106739783
```

Voorbeeld 2

Gemiddeld worden 1 op 2000 huizen per jaar door brand vernield. Bereken de kans dat in een gemeente van 6000 huizen er 4 of meer huizen door brand vernield worden.

Uitwerking

$\lambda = 3$

$P(X \geq 4) = 1 - P(X \leq 3) \approx 0{,}3528$

$GR: 1 - \text{poissoncdf}(3,3)$

Opgave 1

In een textielfabriek worden rollen stof vervaardigd met een lengte van 50 meter per rol. Het aantal weeffouten per rol is Poisson-verdeeld met een bijbehorende verwachtingswaarde van 1 weeffout per rol.

Bij de kwalitatieve keuring van de rollen stof worden deze gescheiden in rollen van “A-kwaliteit” (met 0 of 1 weeffout per rol) en rollen van “B-kwaliteit” (met twee of meer weeffouten per rol).

a. Bereken de kans dat een willekeurige rol de aanduiding “B-kwaliteit” krijgt.
b. De productieomvang per dag is gelijk aan 2000 meter stof. Hoe groot is de kans dat er op een willekeurige dag tenminste 30 rollen met “A-kwaliteit” worden gemaakt?

Opgave 2

Bij een callcenter komen gemiddeld 10 telefoontjes per uur binnen.

a. Bereken de kans dat er in een uur meer dan 15 telefoontjes binnen komen.
b. Bereken de kans dat er in een uur minder dan 5 telefoontjes binnen komen.

Het komt wel ’s voor dat er gedurende een kwartier geen telefoontjes binnen komen.

c. Bereken de kans dat er een kwartier lang geen telefoontjes binnen komen.

Hoofdstuk 3 - waargenomen en berekende waarden

Gegeven is de volgende tabel:

Tabel 1 Aantal dodelijke ongelukken veroorzaakt door een trap van een paard van 10 Pruisische legerkorpsen in een periode van 20 jaar. (1875-1894) (L.v.Bortkiewicz, Das Gesetz der kleinen Zahlen, Leipzig, 1898)		
	Aantal jaren met x doden per korps	
x	**Gemeten**	**Berekend**
0	109	109
1	65	66
2	22	20
3	3	4
4	1	1
>5	0	0

Om de waarde van λ vast te stellen wordt eerst met behulp van de tabel hiernaast het totaal aantal ongelukken met dodelijke afloop berekend, dat is 122 (ga na!).

Het totaal aantal jaren is 200. Dus de kans op een ongelukje met fatale afloop is 0,61.

Dus λ = 0,61 doden per jaar.

We gaan er van uit dat we hier te maken hebben met een Poissonverdeling.

Met de Poisson-verdeling kan je dan de kansen berekenen met:

$$P(X = k) = e^{-0,61} \cdot \frac{0,61^k}{k!}$$

Zie de tabel hiernaast.

Vermenigvuldigen van de rechter kolom in tabel 2 met 200 levert de rechter kolom op in tabel 1.

Tabel 2 Berekende kansen	
k	**P(X=k)**
0	0,543
1	0,331
2	0,101
3	0,021
4	0,003
>5	0,000

Opgave 3

In een medische hulpcentrum heeft men bijgehouden hoeveel cliënten zich in een weekend melden voor 'eerste hulp bij ongelukken'.

- Bereken de theoretische aantallen in het geval je hier te maken zou hebben met een Poissonverdeling.
- In hoeverre komen beide frequenties overeen?

aantal	frequentie
0	0
1	2
2	8
3	12
4	6
5	3
6	2

Hoofdstuk 4 - verjaardagsproblemen

Bij het afleiden van de Poissonverdeling zijn we uitgegaan van een speciaal geval van de binomiale verdeling. Bij de binomiale verdeling ga je er van uit dat de verschillende uitvoeringen van het kansexperiment onafhankelijk van elkaar zijn. Verrassend genoeg blijkt de Poissonverdeling ook nog heel bruikbaar bij kansproblemen waarbij de verschillende experimenten niet helemaal onafhankelijk van elkaar zijn.

Opgave 4

Een kroeg heeft 26 stamgasten. Bereken **exact** de kans dat er minimaal twee gasten op dezelfde dag jarig zijn.

Poisson benadering van het verjaardagsprobleem

Je kunt zo'n kans als bij het verjaardagsprobleem (en andere kansproblemen) ook benaderen met de Poissonverdeling. Je kunt het 'experiment' opvatten als een reeks deelexperimenten. Je kiest steeds een persoon uit de groep en kijkt of er iemand anders is die op dezelfde dag jarig is. De kans dat er iemand op dezelfde dag jarig is is gelijk aan $\frac{1}{365}$. Dat doe je vervolgens voor alle mogelijk tweetallen die je kan maken. Het totaal aantal mogelijke tweetallen in een groep van m personen is gelijk aan $\binom{m}{2}$.

$$\binom{m}{2} = \frac{m!}{2!\cdot(m-2)!} = \frac{1\cdot 2\cdot 3\cdot\ldots\cdot m}{1\cdot 2\cdot 1\cdot 2\cdot 3\cdot\ldots\cdot m-2} = \frac{(m-1)\cdot m}{2} = \frac{1}{2}m(m-1)$$

Dus bij een groep van bijvoorbeeld 26 personen zijn er 325 mogelijke tweetallen. Elk met een kans van $\frac{1}{365}$. De verwachtingswaarde is ongeveer 0,89. Je kunt de kansverdeling dan benaderen met de Poissonverdeling met $\lambda = 0{,}89$.

$$P(X=k) = e^{-0{,}89}\cdot\frac{0{,}89^k}{k!}$$

P(X=0) is dan de kans dat er **niet** een tweetal op dezelfde dag jarig is.

$$P(X=0) = e^{-0{,}89}\cdot\frac{0{,}89^0}{0!} = e^{-0{,}89} \approx 0{,}410$$

De kans dat er minimaal 2 personen op dezelfde dag jarig zijn is ongeveer gelijk aan 0,590. Dat is een opmerkelijk goede benadering als je dat vergelijkt met het **exacte** antwoord van opgave 4.

Maar 't is een benadering. De 'deelexperimenten' zijn niet onafhankelijk van elkaar. We hebben nu gedaan alsof dat wel zo is. Het aantal geboortedagen is echter groot, zodat we die afhankelijkheid 'licht' op kunnen vatten.

De kunst is nu om het verjaardagsprobleem in allerlei andere situaties te herkennen en te ontdekken wat de 'geboortedagen' en wat de 'personen' zijn. Het aantal mogelijke 'geboortedagen' moet wel voldoende groot zijn voor een goede benadering met de Poissonverdeling.

Opgave 5

Stel dat je met je vriend de weddenschap afsluit dat van de eerstvolgende vijftien auto's die langskomen tenminste twee auto's nummerborden met twee gelijke eindcijfers hebben. Wat is de kans dat je wint? Bereken zowel de exacte waarde van de kans als de Poisson benaderingswaarde. [2]

Opgave 6

Wat is de kans dat bij 10 draaiingen van het roulettewiel het balletje twee of meer keer op eenzelfde getal valt? Bij Europees roulette valt het balletje op één van de getallen 0, 1, …, 36. Bereken zowel de exacte waarde als de Poisson benaderingswaarde.[3]

Het bijna-verjaardagsprobleem

Het probleem luidt als volgt[4]:

> "Wat is de kans dat in een willekeurig gevormde groep van m personen twee of meer personen binnen één dag van elkaar jarig zijn?"

Een exacte oplossing is niet zo eenvoudig, maar met de Poisson benadering is het goed te doen. We hebben nog steeds te maken met $\binom{m}{2}$ deelexperimenten. De kans op een deelexperiment met 'succes' is gelijk aan $\frac{3}{365}$. Er geldt: $\lambda = \binom{m}{2} \cdot \frac{3}{365} = \frac{1}{2}m(m-1) \cdot \frac{3}{365}$

Hoofdstuk 5 - het dekpuntenprobleem

Voorbeeld

Op een discofeest gooit iedere mannelijke bezoeker bij binnenkomst zijn baseballpet op een hoop in een hoek en pakt bij vertrek zonder te kijken een pet van de hoop. Wat zou nu de kans zijn dat tenminste één bezoeker met zijn eigen pet weggaat[5]?

Het dekpuntenprobleem

Neem eens aan dat je 10 foto's van presidenten van de Verenigde Staten hebt en 10 naambordjes met de namen. Je hangt de naambordjes willekeurig bij de foto's. Wat zou dan de verwachtingswaarde zijn van het aantal goede namen bij de goede president?

Heel **grof gezegd** zou je kunnen zeggen dat de kans dat een bordje 'goed hangt' gelijk is aan $\frac{1}{10}$. Je voert dit 'experiment' 10 keer uit, dus de verwachtingswaarde is $\frac{1}{10} \cdot 10 = 1$.

Dat klopt natuurlijk niet want de verschillende 'deelexperimenten' zijn niet onafhankelijk, maar als m maar voldoende groot is dan klopt het aardig. Gezien het bovenstaande ligt het voor de hand om dekpuntproblemen te benaderen met de Poissonverdeling met $\lambda = 1$.

[2] Uit Poisson, de Pruisen en de lotto - Epsilon uitgaven
[3] Uit Poisson, de Pruisen en de lotto - Epsilon uitgaven
[4] Uit Poisson, de Pruisen en de lotto - Epsilon uitgaven
[5] Uit Poisson, de Pruisen en de lotto - Epsilon uitgaven

In hoofdstuk 5 van '**Poisson, de Pruisen en de Lotto**' kan je nalezen hoe dat zo werkt. Kennelijk maakt het niet uit of je nu 10 of 10.000 foto's hebt. In het boekje las ik dat vanaf n=10 de exacte waarden van de kansen en de benaderingswaarden met de Poissonverdeling in 8 of meer decimalen overeen blijken te stemmen.

Opgave 7

Een docent heeft een diagnostische toets gemaakt waarbij een leerling bij 10 gegeven vergelijkingen en oplossingen de juiste oplossing bij de juiste vergelijking moet zetten.

a. Als de leerling de oplossingen willekeurig bij de vergelijkingen zet bereken dan de kans in 3 decimalen dat er geen enkele oplossing bij de goede vergelijking staat.
b. Bereken de kans in 3 decimalen op minimaal 3 goed als een leerling de oplossingen willekeurig bij de vergelijkingen zet.
c. Een leerling die de oefening maakt heeft 6 opgaven goed gemaakt. De docent denkt dat deze leerling gegokt heeft. Vind je dat terecht?

Opgave 8

Voor wiskundestudenten wordt eens per jaar een werkwerk (Wisweek) georganiseerd. De organisatie vraagt aan alle tachtig studenten om één cd-tje mee te nemen met hun lievelingsmuziek. Overdag en 's avonds wordt er muziek gedraaid. Maar het is niet verstandig om 's nachts de cd's te laten liggen in de open kantine, dus voor het naar bed gaan pakt iedereen willekeurig een cd om die de volgende dag weer af te geven.

a. Bereken de kans dat geen enkele student zijn eigen cd heeft gepakt.
b. Bereken de kans dat twee of meer studenten hun eigen cd hebben gepakt.

De Wisweek duurt drie dagen en nachten.

c. Bereken de kans dat er tijdens de Wisweek een avond was dat tenminste één student zijn eigen cd heeft gepakt.

EINDE

Literatuur:

- Poisson, de Pruisen en de lotto
  De Poisson verdeling en haar toepassingen
  Henk Tijms, Frank Heierman en Rein Nobel
  Epsilon Uitgaven in samenwerking met de Nederlandse Vereniging van Wiskundeleraren
  Utrecht, 2000
  Zebrareeks

Uitwerkingen van de opgaven.

Opgave 1

a. X: aantal weeffouten per rol
   X~ Poissonverdeeld met $\lambda = 1$
   $P(X \geq 2) = 1 - P(X \leq 1) \approx 1 - 0,736 = 0,264$

b. P(A kwaliteit) = 0,736
   X: aantal rollen met A-kwaliteit
   X~ binomiaal verdeeld met n=40 en p=0,736
   $P(X \geq 30) = 1 - P(X \leq 29) \approx 1 - 0,499 = 0,501$

Opgave 2

a. X:aantal telefoontjes per uur
   X~Poissonverdeeld met $\lambda = 10$
   $P(X > 15) = 1 - P(X \leq 15) \approx 1 - 0,951 = 0,049$

b. X:aantal telefoontjes per uur
   X~Poissonverdeeld met $\lambda = 10$
   $P(X < 5) = P(X \leq 4) \approx 0,029$.

c. Neem $\lambda = \frac{10}{4} = 2\frac{1}{2}$ en bereken $P(X = 0) = 0,082$.

Opgave 3

a.

aantal	frequentie	aantal·freq	berekend	theoretisch
0	0	0	0,0415	1
1	2	2	0,1321	4
2	8	16	0,2101	7
3	12	36	0,2229	7
4	6	24	0,1773	6
5	3	15	0,1128	4
6	2	12	0,0598	2
	33	**105**		
	λ=	3,1818		

b.

De frequenties komen aardig overeen, met uitzondering van k=3. Voor de rest lijkt het aardig te kloppen.

Opgave 4

Bereken de kans dat iedereen op een andere datum jarig is en neem de complementaire kans[6].

```
365 nPr 26/365^2
6
            .4017591799
```

$$P(\text{niemand}) = \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdot \ldots \cdot \frac{340}{365}$$

$$P(\text{niemand}) = \frac{(365)_{26}}{365^{26}} \approx 0{,}4018$$

$$P(\text{minimaal 2 op 1 dag}) \approx 1 - 0{,}4018 = 0{,}5982$$

Opgave 5

Exact:

$$P(\text{allemaal verschillend}) = \frac{(100)_{15}}{100^{15}} \approx 0{,}331$$

$$P(\text{minstens twee dezelfde}) \approx 1 - 0{,}331 = 0{,}669$$

Benaderd met de Poissonverdeling:

$$\lambda = \frac{1}{2} m(m-1) \text{ en } m = 15 \Rightarrow \lambda = 1{,}05$$

$$P(X=0) = e^{-1{,}05} \cdot \frac{1{,}05^0}{0!} = e^{-1{,}05} \approx 0{,}350$$

Gevraagde kans is 0,650

Opgave 6

$$P(\text{minstens 2 dezelfde}) = 1 - P(\text{allemaal verschillend})$$

$$P(\text{allemaal verschillend}) = \frac{(37)_{10}}{37^{10}} \approx 0{,}263$$

$$P(\text{minstens 2 dezelfde}) \approx 0{,}737$$

$$\lambda = \frac{\frac{1}{2} m(m-1)}{n} = \frac{\frac{1}{2} \cdot 10 \cdot 9}{37} = 1{,}216$$

$$P(\text{minstens 2 dezelfde}) = 1 - e^{-1.216} \approx 0{,}704$$

Opgave 8

Dit is een dekpuntenprobleem, dus $\lambda = 1$.

We gaan benaderen met de Poissonverdeling.

a.

X: aantal goed

$X \sim$ Poissonverdeeld met $\lambda = 1$

$P(X=0) = 0{,}368$

---

[6] Exact? Hier$\rightarrow$ $\dfrac{4368955574049368809099612611947647331684681345038990016381 81}{7303004788367008989935731733484508298905132521247863769531 25}$

b.

$X: \text{aantal goed}$

$X \sim \text{Poissonverdeeld met } \lambda = 1$

$P(X \geq 3) = 1 - P(X \leq 2) \approx 1 - 0{,}920 = 0{,}080$

c.

$H_0 : \mu = 1$

$H_1 : \mu > 1$

Onder H_0 geldt: $P(X \geq 6) = 1 - P(X \leq 5) \approx 1 - 0{,}999 = 0{,}001$

Dit is **veel** kleiner dan 0,05. We verwerpen H_0 en nemen H_1 aan. De leerling heeft (waarschijnlijk) niet gegokt. De docent heeft ongelijk.

Opgave 9

a. Binomiale verdeling met p = 1/80 en n = 80, dus $\mu = np = 1$
We gaan de Poissonverdeling gebruiken met $\lambda = 1$
$P(X = 0) = e^{-1} = 0{,}3678794 \approx 0{,}3679$

b. $P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 0{,}2642411 \approx 0{,}2642$

c. Dit is een nieuwe binomiale verdeling met n = 3 en als alternatief:
A = niemand heeft zijn eigen cd gepakt met kans $p = e^{-1} \approx 0{,}3679$ (zie vraag a)
B= één of meer studenten hebben hun eigen cd met kans q = 1 - p
Stochast Y telt het aantal avonden dat niemand zijn eigen cd pakt
P(een avond, tenminste..) = $P(Y < 3) = 1 - P(Y = 3) = 1 - (e^{-1})^3 = 1 - 0{,}049781 =$
$0{,}9502129 \approx 0{,}9502$