Waarskynlikhede 3 –Kontinue Modelle

In ’n vorige artikel[1] is u aan verskeie diskrete waarskynlikheidsmassafunksies oftewel diskrete waarskynlikheidsverdelings voorgestel. Daar is nog verskeie ander wat ook baie nuttig is. Volledigheidshalwe gee ek die bekendes en voeg nog ’n paar by. Hoe hierdie waarskynlikheidsverdelings onstaan het en die besondere voorwaardes en gebruike daarvoor is in die genoemde artikel, sowel as nog ’n vorige artikel uiteengesit.[2] Daarna net so een en ander oor die verdelings vir kontinue data.

Uniforme diskrete waarskynlikheidsverdeling:

$$P(X=x)=f(x)=\frac{1}{n} \quad x=1,2,3,\dots,n$$
$$=0 \quad \text{vir enige ander waarde van } x$$

Binomiaal waarskynlikheidsverdeling:

$$P(X=x)=f(x)=\binom{n}{x}p^{x}(1-p)^{n-x} \qquad x=0,1,2,3,\dots,n$$
$$=0 \qquad \text{andersins}$$

Hipergeometriese waarskynlikheidsverdeling:

$$P(X=x)=f(x)=\begin{cases}\dfrac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}} & x=0,1,\dots,\min(K;n)\\ 0 & \text{andersins}\end{cases}$$

Bernoulli-waarskynlikheidsverdeling:

$$P(X=x)=f(x)=p^{1-x}(1-p)^{x} \quad x=0 \text{ of } 1, \text{ ook geskryf as } x=0,1$$
$$=0 \quad \text{vir enige ander waarde van } x$$

Poisson-waarskynlikheidsverdeling:

$$g(x)=\frac{(\lambda t)^{x}e^{-\lambda t}}{x!}, \quad x=0,1,2,3,\dots$$
$$=0 \qquad \text{andersins}$$

Ons voeg twee nuwe diskrete waarskynlikheidsverdelings by, nl. die

- Negatief Binomiaal waarskynlikheidsverdeling (Pascalverdeling) en die

[1] https://www.litnet.co.za/wp-content/uploads/2020/07/2020-06-30-Waarskynlikhede-2-%E2%80%93-Diskrete-Modelle.pdf
[2] https://www.litnet.co.za/wp-content/uploads/2020/06/2020-06-27-Teltegnieke-en-waarskynlikhede-1.pdf

- Meetkundige (geometriese) waarskynlikheidsverdeling.

Een baie belangrike opmerking: Maak seker dat u presies weet waarvoor die veranderlike in elke model staan, nl. tyd of aantal of wat ook al, anders is u gedoem tot verwarring.

Die vraag wat ons met die negatief binomiaal waarskynlikheidsverdeling wil beantwoord is: Wat is die waarskynlikheid dat die $r$-ste sukses met die $x$-ste eksperiment behaal word? Die parameter $r$ is die gespesifiseerde aantal suksesse wat ek wil verkry en die veranderlike $X$ is die aantal eksperimente daarvoor benodig. Let op die verskil tussen hierdie verdeling en die binomiaalverdeling. Die waarskynlikheid vir 'n sukses bly vas, nl. $p$ , maar die binomiaalverdeling vra: Wat is die waarskynlikheid van $X$ suksesse in 'n vaste aantal herhalings $n$ ? By die negatief binomiaal waarskynlikheidsverdeling is die aantal eksperimente benodig, nl. $X$ die veranderlike.

Beskou $X$ as die aantal eksperimente tot en met die gespesifiseerde $r$-ste sukses. Die waarskynlikheid $(1-p)$ van 'n mislukking by elke eksperiment is natuurlik ook vas, net soos die waarskynlikheid van 'n sukses per eksperiment ook vas is. Dit vereis $(r-1)$ suksesse tot en met die $(x-1)$-ste eksperiment ongeag die volgorde van suksesse en mislukkings, d.w.s. 'n binomiaal situasie met 'n direk daaropvolgende sukses by die $x$-ste eksperiment of herhaling. Dus eerstens:

$$\begin{aligned}
&P\left[(r-1) \text{ suksesse in } (x-1) \text{ eksperimente}\right] \\
&=\binom{x-1}{r-1} p^{r-1}(1-p)^{(x-1)-(r-1)} \\
&=\binom{x-1}{r-1} p^{r-1}(1-p)^{x-r}.
\end{aligned}$$

Tweedens dan die waarskynlikheid vir die $r$-ste sukses met die $x$-ste eksperiment bygewerk:

$$\begin{aligned}
&P\left[\ (r-1) \text{ suksesse in } (x-1) \text{ eksperimente} \cap \text{ nog een eksperiment wat suksesvol is}\right] \\
&= P\big((r-1) \text{ suksesse in } (x-1) \text{ eksperimente}\big) \cdot P(\text{nog een eksperiment wat suksesvol is}) \\
&=\binom{x-1}{r-1} p^{r-1}(1-p)^{x-r} \cdot p \\
&=\binom{x-1}{r-1} p^{r}(1-p)^{x-r}
\end{aligned}$$

Dus is die funksie ter sprake

$$\begin{aligned}
P(X=x)=f(x)&=\binom{x-1}{r-1} p^{r}(1-p)^{x-r}, \qquad x=r,r+1,r+2,\ldots \\
&=0 \qquad\qquad \text{vir enige ander waarde van } x
\end{aligned}$$

Die meetkundige waarskynlikheidsverdeling is 'n spesiale geval van die negatief binomiaal waarskynlikheidsverdeling met $r = 1$, d.w.s. dit gee vir ons die waarskynlikheid dat die eerste sukses met die $X$-ste ekperiment verkry sal word. Dus lyk die funksie soos volg:

$$f(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, \ldots \\ 0 & \text{andersins} \end{cases}$$

Soms word $X$ gedefinieer as die aantal eksperimente vóór die eerste sukses. Dit lei tot 'n klein wysiging in die funksie soos volg, met 'n logiese wysiging in die definisiegebied, oftewel die populasie van $X$:

$$f(x) = \begin{cases} p(1-p)^{x} & x = 0, 1, 2, \ldots \\ 0 & \text{andersins} \end{cases}$$

Voorbeeld (betroubaarheid, *reliability*)[3]:

Drie identiese onafhanklike rekenaars se betroubaarheid word ondersoek. Een word gebruik en twee staan in as rugsteun. Die rekenaars is onafhanklik van mekaar. Die waarskynlikheid vir faling van 'n rekenaar op enige gegewe stadium as dit gebruik word is 0.0005. Die rekenaars is nie herstelbaar nie. Spesifiseer die veranderlike as $X$, die aantal kere van gebruik totdat al drie die rekenaars, en dus die stelsel, gefaal het. Nou word die aantal kere van gebruik vir elk van die drie rekenaars tot faling aangedui met $x_1$, $x_2$ en $x_3$ sodat die aantal kere van gebruik tot finale faling aangedui word deur

$$X = x_1 + x_2 + x_3, \text{ wat beteken dat}$$

$$X \sim neg\,bin(p = 0.0005; r = 3)$$

$$P(X = x) = f(x) = \binom{x-1}{3-1} 0.0005^3 \cdot 0.9995^{x-3} \qquad x = 3, 4, 5, \ldots$$

Nou kan ons die waarskynlikheid bereken dat die rekenaarstelsel binne 5 geleenthede van gebruik sal faal, nl. met die derde gebruik of vierde gebruik of vyfde gebruik. Let op dat die stelsel nie met die eerste of tweede gebruik kan faal nie, want al sou 'n rekenaar faal, is daar steeds een van die rugsteun-rekenaars beskikbaar. Eers as die derde rekenaar faal, faal die stelsel. Dus is

$$\begin{aligned} P(X \le 5) &= P(X = 3) + P(X = 4) + P(X = 5) = f(3) + f(4) + f(5) \\ &= \binom{3-1}{3-1} 0.0005^3 \cdot 0.9995^{3-3} + \binom{4-1}{3-1} 0.0005^3 \cdot 0.9995^{4-3} \\ &\quad + \binom{5-1}{3-1} 0.0005^3 \cdot 0.9995^{5-3} \\ &= 1.25 \cdot 10^{-10} + 3.74813 \cdot 10^{-10} + 7.4925 \cdot 10^{-10} \\ &= 1.24906 \cdot 10^{-9} \end{aligned}$$

[3] Montgomery D.C., Runger G.C. (2011). *Applied statistics and probability for engineers,* 5th ed. Wiley.

Daarteenoor kan ons ook die waarskynlikheid dat byvoorbeeld rekenaar nommer een eers by die 5000-ste gebruiksgeleentheid sal faal bereken. In hierdie geval is $X$ die aantal gebruiksgeleenthede tot faling. Dus

$$P(X=x)=f(x)=\begin{cases}0.0005(1-0.0005)^{x-1} & x=1,2,\ldots \\ 0 & \text{andersins / } \textit{otherwise}\end{cases}$$

Sodat

$$\begin{aligned}\therefore\ P(X=5000)=f(5000)&=0.0005(1-0.0005)^{4999}\\ &=4.10168\cdot 10^{-5}\end{aligned}$$

Diskrete modelle kom in baie besondere omstandighede voor en het dikwels sogenaamde kontinue ekwivalente. Ons gaan net na ’n paar basiese beginsels van die kontinue modelle kyk en dit dan daar laat. Die rede hiervoor is dat ’n mens redelik vaardig moet wees in integrasietegnieke om dit met gemak te kan hanteer. Dit verhoed ons egter nie om ’n goeie idee te vorm van hoe ’n mens daarmee te werk kan gaan nie.

Ons gaan slegs enkele kontinue modelle noem. Die benaming is natuurlik verkeerd, want dis nie die model wat diskreet of kontinu is nie, maar dis modelle wat geld wanneer die veranderlike diskreet of kontinu van aard is. Dit is egter konvensie om die benamings so te gee.

Een van die kontinue modele wat ons bespreek is sekerlik een van die vernaamste in die statitiek, nl. die normaal waarskynlikheidsverdeling wat ’n waarskynlikheids**digtheid**s-funksie het, teenoor die waarskynlikheids**massa**funksie van die diskrete verdelings. Die normaalverdeling kom byvoorbeeld voor waar die veranderlike enige waarde op die reële as kan aanneem. Die feit dat die punte op ’n reële lyn dig is en nie diskreet nie, lei tot die benaming van digtheidsfunksie.

Baie belangrik: Waar $f(x)$ in die diskrete modelle op waarskynlikhede dui, is dit nie die geval by kontinue modelle nie. Hoeveel punte is op ’n lynstuk van lengte een? Ons weet dis oneindig. Dus is die waarskynlikheid vir enige gespesifiseerde punt $P(X=x)=\frac{1}{\infty}$ wat vir alle praktiese doeleindes nul is. Dit beteken nie dis onmoontlik nie, maar die waarskynlikheid daarvoor is onberekenbaar klein en ons hanteer dit as ’n nul. In die kontinue geval kan ons dus slegs die waarskynlikheid van ’n suksesvolle gebeurtenis oor ’n interval bepaal. Gebruik byvoorbeeld die geslote interval $[1,4]$ as die definisiegebied van $X$ of dan die populasie van moontlike waardes van $X$. Dan is

$$P(1.5\le X\le 3.5)=\frac{3.5-1.5}{4-1}=\frac{2}{3}.$$

Dit lyk logies, nè? U het onbewustelik hier van ’n verskeidenheid aannames gebruik gemaak. So byvoorbeeld kon ons hierdie uitdrukking geskryf het as

$$
\begin{aligned}
P(1.5 \le X \le 3.5) &= P(X = 1.5) + P(1.5 < X < 3.5) + P(X = 3.5) \\
&= 0 + P(1.5 < X < 3.5) + 0 \\
&= \frac{2}{3}
\end{aligned}
$$

Waar die <-teken teenoor die $\le$-teken by diskrete verdelings kritiek is, is dit by kontinue veranderlikes vir die standaard berekenings in die algemeen van weinig belang.

’n Verdere eienskap wat u gebruik het is dié van die kontinue uniforme waarskynlikheids-verdeling met digtheidsfunksie:

$$
\begin{aligned}
f(x) &= \frac{1}{b-a} \qquad a \le x \le b \\
&= 0 \qquad \text{andersins}
\end{aligned}
$$

In hierdie geval is $a = 1$ en $b = 4$.

Die ooreenkomstige verdelingsfunksie, ook genoem die kumulatiewe verdeling is

$$
\begin{aligned}
F(x) &= \frac{x-a}{b-a} \qquad a \le x \le b \\
&= 0 \qquad \text{andersins}
\end{aligned}
$$

oftewel

$$
\begin{aligned}
F(x) &= \frac{x-1}{4-1} = \frac{x-1}{3} \qquad 1 \le x \le 4 \\
&= 0 \qquad \text{andersins}
\end{aligned}
$$

In die voorbeeld het ons dus

$$
\begin{aligned}
P(1.5 \le X \le 3.5) &= P(X \le 3.5) - P(X \le 1.5) \\
&= F(3.5) - F(1.5) \\
&= \frac{3.5-1}{3} - \frac{1.5-1}{3} \\
&= \frac{2}{3}
\end{aligned}
$$

Nog ’n manier van bereken is met behulp van integrasie soos volg:

$$P(1.5 \le X \le 3.5) = \int_{1.5}^{3.5} \frac{1}{3}\,dx$$
$$= \left[\frac{1}{3}x\right]_{1.5}^{3.5}$$
$$= \frac{1}{3}[3.5 - 1.5]$$
$$= \frac{2}{3}$$

U werk hier met 'n figuur tussen die waardes 1 en 4 op die $x$-as wat die onderste sy van 'n reghoek met 'n hoogte van $f(x) = \frac{1}{b-a} = \frac{1}{3}$ vorm. Die totale area onder die kurwe/reguit lyn $f(x) = \frac{1}{b-a} = \frac{1}{3}$ is gelyk aan 1, nl. die totale waarskynlikheid van die gedefinieerde populasie.

'n Terloopse opmerking is ter sake. Wat beteken die uitdrukking "$f(x) = 0$ andersins" wat elke keer by 'n massafunksie of digtheidsfunksie gegee word? 'n Digtheidsfunksie, soos 'n massafunksie, moet te alle tye vir alle waardes van $X$ gedefinieer wees. Die kontinue uniforme digtheidsfunksie lyk in der waarheid soos volg:

$$\begin{aligned} f(x) &= 0 && x < a \\ &= \frac{1}{b-a} && a \le x \le b \\ &= 0 && x > b \end{aligned}$$

sodat $X$ oor die hele interval van $-\infty$ tot $\infty$ gedefinieer is.

Die normaalverdeling se digtheidfunksie het o.a. uit 'n studie van afwykings van die norm in produksieprosesse ontstaan. Hierdie afwykings is benader deur die volgende funksie op data te pas waar $x$ op die afwykings dui

$$erf(x) = e^{-\frac{1}{2}\frac{(x-a)^2}{b}} \quad \text{vir } \infty < x < \infty, \quad b > 0$$

Die notasie *erf*() staan vir *error function*. Die totale oppervlakte onder 'n waarskynlikheids-digtheidsfunksie moet natuurlik gelyk wees aan een om vir waarskynlikheidberekenings geskik te wees, sodat die funksie aangepas moet word deur dit te standaardiseer na een. Die integraal van *erf*($x$) oor die definisiegebied van $X$ vanaf $-\infty$ tot $\infty$ is $\sqrt{2\pi b}$. Al wat ons dus te doen staan is om deur hierdie faktor te deel en die saak is reg. Verder het die konvensie ontstaan om $a$ en $b$ deur $\mu$ en $\sigma^2$, die verwagte waarde of gemiddelde en die variansie van die normaalverdeling te vervang – wat nie te moeilik is om te bewys nie. Die uiteindelike normaal digtheidsfunksie lyk dan soos volg:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \frac{-(x-\mu)^2}{2\sigma^2}, \quad \infty < x < \infty$$

Die variansie is $\sigma^2$ en die vierkantswortel hiervan, $\sigma$ staan bekend as die standaardafwyking. ’n Tipiese voorbeeld van ’n toepassing van die normaalverdeling is die voorkoms van die IK’s van ’n redelike homogene groep mense. Veronderstel ons het ’n groep mense waarvan die gemiddelde IK $\mu = 117$ is met ’n variansie van $\sigma^2 = 100$. Die standaard skryfwyse hiervoor is $X \sim N(117,100)$ om ooreen te stem met die konvensie van $X \sim N(\mu,\sigma^2)$. Die waarskynlikheid dat ’n ewekansig gekose persoon ’n IK tussen 107 en 127 sal hê is nou

$$
\begin{aligned}
P(107 < X < 127) &= \int_{107}^{127} \frac{1}{\sqrt{2\pi}10} \frac{-(x-117)^2}{2\cdot 100} dx \\
&= 0.8413 - 0.1586 \\
&= 0.6827
\end{aligned}
$$

Hierdie integraal kan nie analities bereken word nie en ons maak van rekenaarsagteware (voorheen tabelle) gebruik om die antwoord te kry. Baie sakrekenaars het ook al vandag hierdie fasiliteit. Sommige sakrekenaars vereis dat die integrasiegrense genormaliseer moet word, sodat die integraal dan soos volg lyk:

$$
P(107 < X < 127) = \int_{107}^{127} \frac{1}{\sqrt{2\pi}10} \frac{-(x-117)^2}{2\cdot 100} dx
$$

$$
\text{Stel } z = \frac{x-\mu}{\sigma} = \frac{x-117}{10}, \ \Rightarrow x = 117 + 10z
$$

$$
\text{sodat } \frac{dx}{dz} = 10, \ \text{ of } dx = 10dz. \text{ Substitueer, dan is}
$$

$$
\begin{aligned}
P(107 < X < 127) &= \int_{-1}^{1} \frac{1}{\sqrt{2\pi}} \frac{-z^2}{2} dz \\
&= 0.8413 - 0.1586 \\
&= 0.6827
\end{aligned}
$$

Hier word dus gebruik gemaak van die gestandaardiseerde normaalverdeling wat soos volg aangedui word: $Z \sim N(0,1)$. Hier word ’n belangrike eienskap van die normaalverdeling geïllustreer, nl. dat 68.27% van ’n normaal verdeelde populasie se waardes binne een standaardafwyking van sy gemiddelde af is.

Die normaalverdeling is buitengewoon nuttig. Indien steekproewe groot is, neig menige verdelings om hulle soos die normaalverdeling te gedra. (Statistici tel maar sleg. Groot is ’n baie relatiewe begrip en 30 is dikwels al “groot genoeg”.) Indien jy met ’n binomiaalverdeling werk en $np > 10$ is jy al redelik veilig wat akkuraatheid betref deur dit as ’n $N(np, np(1-p))$-verdeling te hanteer. Daar is wel sogenaamde kontinuïtetskorreksies ter sprake, maar dis slegs tegnies en in die praktyk nie van groot belang nie. Vele ander diskrete en kontinue modelle streef na die normaalverdeling indien *n* groot genoeg is. Ek gaan egter nie nou verder op hierdie aspek in nie, want dan begeef ons ons baie gou op die veld van die sg. steekproefverdelings, ’n ander geweldige groot en belangrike veld in die studie van statistiek, maar te lywig om hier aan te pak.

Vroeër is genoem dat verskeie diskrete verdelings ekwivalente in die kontinue verdelings het. Die geometriese verdeling en negatief binomiaalverdelings is twee daarvan, maar daar is nog meer.

Die geometriese verdeling beskryf die **aantal** eksperimente voor die eerste sukses. Die eksponensiële verdeling beskryf die **wagtyd** tot die eerste sukses en dit lyk soos volg:

$$\begin{aligned} f(t) &= \lambda e^{-\lambda t}, \quad t \geq 0 \\ &= 0 \qquad \text{andersins} \end{aligned}$$

waar $\lambda$ die gemiddelde voorkomskoers van 'n gebeurtenis per 'n gedefinieerde tydseenheid is. Voorbeeld: Telefoonoproepe daag teen 'n gemiddelde koers van 15 oproepe per uur by 'n sentrale op. Wat is die waarskynlikheid dat die volgende oproep nie binne die volgende 5 minute sal opdaag nie? Die vraag is dus: Bereken $P\left(T > 5 \text{ vir die eerste oproep} \middle| \lambda = 15 \text{ per uur}\right)$. Die antwoord kan nou bereken word deur gebruik te maak van minuut eenhede óf 5 minute eenhede óf uur eenhede. Dit maak geen verskil aan die antwoord nie.

$$\begin{aligned} P(T > 5\,\text{min}) = P\left(\text{wagtyd} > \frac{1}{12} \text{ uur}\right) &= \int_{\frac{1}{12}}^{\infty} 15 \cdot e^{-15t} dt, \qquad \text{(dus in uur eenhede)} \\ &= \left.\frac{15e^{-15t}}{-15}\right|_{\frac{1}{12}}^{\infty} = -\left(e^{-15(\infty)} - e^{-15\left(\frac{1}{12}\right)}\right) \\ &= -(0 - 0.2865) \\ &= 0.2865 \end{aligned}$$

Die negatief binomiaalverdeling vra: Indien $\lambda$ die gemiddelde voorkomskoers van 'n gebeurtenis per 'n gedefinieerde tydseenheid is wat is die waarskynlikheid is dat ek die *r*-ste sukses eers met die $X$ -ste **herhaling** sal kry. Die gammaverdeling, ook bekend as die Erlang-verdeling, vra wat die waarskynlikheid is dat ek minstens $T$ **tydseenhede moet wag** vir die *r*-ste sukses en dit lyk soos volg:

$$\begin{aligned} f(t) &= \frac{\lambda^r t^{r-1} e^{-\lambda t}}{\Gamma(r)} \quad \text{waar } \Gamma(r) = (r-1)! \text{ en t} > 0 \\ &= 0 \quad \text{andersins} \end{aligned}$$

Hierdie uitdrukking hou natuurlik direk verband met Euler se gamma-integraal,[4] nl.

$$\int_0^{\infty} \lambda^r t^{r-1} e^{-\lambda t} dt = \Gamma(r) = (r-1)!$$

[4] Die bewys vir Euler se integraal sal by 'n latere geleentheid gegee word.

Let op dat as $r=1$ volg dit dat $f\left(t\right)=\frac{\lambda^{r}t^{r-1}e^{-\lambda t}}{\Gamma\left(r\right)}=\frac{\lambda^{1}t^{1-1}e^{-\lambda t}}{\Gamma\left(1\right)}=\lambda e^{-\lambda t}$ wat natuurlik niks anders is as die eksponensiële waarskynlikheidsdigtheidsfunksie hierbo nie.

Voorbeeld: Die gemiddelde tussen-aankomstyd van foute op 'n produksielyn is 30 minute, d.w.s. gemiddeld een fout per 30 minute. Die betrokke veranderlike is eksponensiaal verdeel. Wat is die waarskynlikheid dat minstens 90 minute sal verloop **voordat** die derde fout voorkom, d.w.s. die waarskynlikheid dat ek meer as 90 minute moet wag vir derde sukses? Ek gebruik hier minuut eenhede vir tyd.

$$
\begin{aligned}
P\left(T>90\right)&=\int_{90}^{\infty}\frac{\left(\frac{1}{30}\right)^{3}t^{3-1}\exp\left(-\frac{1}{30}t\right)}{\Gamma\left(3\right)}dt \qquad \text{Stel y}=\frac{1}{30}t\therefore dt=30dy\\
&=\int_{y=\frac{90}{30}=3}^{y=\frac{\infty}{30}=\infty}\frac{\left(\frac{1}{30}\right)^{3}\left(30y\right)^{3-1}\exp\left(-\frac{1}{30}\left(30y\right)\right)}{\Gamma\left(3\right)}30dy\\
&=\int_{3}^{\infty}\frac{y^{3-1}\exp\left(-y\right)}{\Gamma\left(3\right)}dy\\
&=1-0.577=0.423
\end{aligned}
$$

Bostaande integraal is lastig om op te los, maar weereens is rekenaarsagteware (of tabelle) handig. Daar is egter 'n ander en selfs makliker manier om na hierdie probleem te kyk. Die vraag sou soos volg gestel kon wees: Wat is die waarskynlikheid van **minder as** drie foute in die eerste 90 minute. Dan volg u 'n Poisson-verdeling[5] benadering en die berekening lyk soos volg met $\lambda t=3$ per 90 minute:

$$
\begin{aligned}
P\left(X\ \le\ 2\right)\ &=\sum_{x=0}^{2}\frac{e^{-\lambda t}\left(\lambda t\right)^{x}}{x!}\quad \text{met } \lambda t=\frac{1}{30}\cdot 90=3\\
&=P\left(0\right)\ +\ P\left(1\right)\ +\ P\left(2\right)\\
&=\frac{e^{-3}\left(3\right)^{0}}{0!}+\frac{e^{-3}\left(3\right)^{1}}{1!}+\frac{e^{-3}\left(3\right)^{2}}{2!}\\
&=e^{-1.5}\left(\frac{1}{0!}+\frac{3}{1!}+\frac{9}{2!}\right)\\
&=\ 0,423
\end{aligned}
$$

Soms soek ons 'n waarskynlikheidsverdeling waar die veranderlike slegs positief kan wees, maar met sekere van die eienskappe van die normaalverdeling. Hier dink ons aan alternatiewe soos die lognormaal- en gammaverdelings. Ander kontinue waarskynlikheidsverdelings kan gemanipuleer word om oor enige interval wat toepaslik is te geld. Hier is die beta en Weibull-waarskynlikheidsmodelle baie nuttig. Daar is 'n magdom sulke modelle. Verder vind ons

[5] https://www.litnet.co.za/wp-content/uploads/2020/07/2020-06-30-Waarskynlikhede-2-%E2%80%93-Diskrete-Modelle.pdf

dikwels dat 'n empiriese verdeling goed deur 'n ad hoc-verdeling sonder spesifieke verband tussen die fisiese waarnemings (en hulle oorsprong) en die parameters beskryf kan word.

Waarskynlikheidsmodelle is een van die hoekstene van statistiek en vorm ’n studieveld van sy eie. As u hierin belang stel, nooi ek u uit om dit op die internet te gaan ondersoek. U sal nie teleurgesteld wees nie.[6] Een so ’n skakel sluit ’n geprogammeerde stelsel vir selfleer in.[7] Geniet dit.

---

[6] https://en.wikipedia.org/wiki/List_of_probability_distributions

[7] https://courses.lumenlearning.com/waymakercollegealgebra/chapter/construct-probability-models/