Huiswerk Stochastiek
Week 12

Joris van der Hijden

7 februari 2017

Week 12

Opgave 1

Stel je voor dat je een spel speelt waarbij je 6 euro krijgt als je een zes rolt, maar 1 euro moet bepalen als het geen 6 is en je wilt een idee krijgen van de kans dat je verlies maakt als je honderd keer speelt, dan heb je veel meer aan een schatting om te zien dat de kans klein is.
Stel je daarentegen voor dat je een spoornet beheert en je de kans wilt weten dat een trein ontspoort, dan heb je niet veel aan een schatting. Het zou namelijk kunnen zijn dat er een hele grote spreiding is en dan zou dus al veel eerder een ongeluk plaatsvinden dan je verwacht. In dat geval is een bovengrens veel handiger.

Opgave 2

i. Onder de aanname dat de kans dat treinen onafhankelijk is, weten we dat geldt dat

$$\frac{1}{n}\left(X_1 + X_2 + X_3 + X_4 ... X_n\right) \to \mu.$$

Als we elke trein als een Bernoulli verdeling zien met kans p op te laat komen, heeft elke stochast $\mu = p$.

$$\begin{aligned}(X_1 + X_2 + X_3 + X_4 ... X_n) &= \frac{1}{n}\mu \\ \mu &= \frac{1}{100} \cdot 30 = 0.3\end{aligned}$$

ii. Als we aannemen dat de treinen onafhankelijk zijn en verdeeld zijn met een bernoulli verdeling met parameter p.Kunnen we de een bovengrens geven van de kans dat de trein van Lotte meer dat dertig keer te laat is, gegeven dat

$P \leq 25$.

$$\mathbb{P}\left(\sum_{i=1}^{100} X_i \geq 30\right) = \mathbb{P}\left(\sum_{i=1}^{100}(X_i - 100p) \geq 30 - 100 * p\right) = \mathbb{P}\left(\frac{1}{10}\sum_{i=1}^{100}\frac{X_i - 100p}{\sqrt{p(1-p)}} \geq \frac{3-100p}{\sqrt{(1-p)p}}\right)$$

De parameter p is $P/100$ omdat we van een percentage naar een komma getal omrekenen. We willen een overschatting voor de kans, dus hetgene na de ongelijkheid moet zo klein mogelijk worden, want het is positief en we doen straks 1- de kans. op het interval $(0, 25/100]$ is de functie na de ongelijkheid dalend, dus we vullen een zo groot mogelijke p in. Daaruit kunnen we concluderen dat

$$\mathbb{P}\left(\sum_{i=1}^{100} X_i \geq 30\right) < 1 - \frac{1}{2\pi}\int_{-\infty}^{\frac{3-0.1p}{\sqrt{(1-p)p}}} e^{-0.5s^2} ds = 0.13$$

Hetzelfde kunnen we doen met de tweede vraag, we moeten alleen de gelijkheid omklappen en een andere waarde voor p invullen. Omdat we nu niet 1-de kans doen willen we juist een zo groot mogelijke waarde gebruiken. Dat geeft ons dat

$$\mathbb{P}\left(\sum_{i=1}^{100} X_i \leq 30\right) < \frac{1}{2\pi}\int_{-\infty}^{\frac{3-0.1p}{\sqrt{(1-p)p}}} e^{-0.5s^2} ds = 0.15$$

Dit geeft een goede indicatie dat $P \in [25, 35]$ omdat we zo weten dat gegeven dat gegeven $P < 25$ of $P > 35$ de kans klein is dat een steekproef 30 oplevert (de kans op groter en of kleiner gelijk dertig is altijd groter dan de kans op 30[1])

iii. In de spits lopen treinen vaker vertraging op, dus de waarde voor P in de spits ligt waarschijnlijk groter dan nu aangenomen.

Opgave 3

i. In het plaatje is te zien dat als n groter wordt, de binomiale verdeling, gecorrigeerd voor de μ en σ en steeds dichter naar de normale verdeling nadert.

ii. De variantie van een students t verdeelde stochast met parameter ν is

$$Var(X) = \begin{cases} \text{ongedefinieerd} & \text{als } \nu \leq 1 \\ \infty & \text{als } 1 < \nu \leq 2 \\ \frac{\nu}{\nu-2} & \text{als } \nu > 2 \end{cases}$$

iii. De student's t verdeling convergeert veel langzamer dan de binomiale verdeling. Bij de som van 1000 stochasten lijkt de student's t verdeling nog

[1] De kans op precies 30 is natuurlijk nul, maar we kijken naar een interval rond 30

niet echt op. Dit komt waarschijnlijk door dat de variantie zo groot is, namelijk 21.Dan is de spreiding in de waarden groter dus convergeert het langzamer.