Rapport Projekt 1

Kurskod: IX1501
Datum: 2017-11-14

Evan Saboo, saboo@kth.se
Max Kufa, mkufa@kth.se

Uppgift 1: Att vinna en Teddy

Sammanfattning

Uppgift

Den första uppgiften består av fyra deluppgifter som inkluderar ett tärningsspel med chansen att vinna en Teddy.
Tärningsspelet består av fem olika tärningar (d4, d6, d8, d12 och d20 tärningar). Om summan av alla tärningar, X, har ett värde under 10 eller över 45, så vinner man Teddyn.

De fyra deluppgifterna är följande:

- Finn den exakta probabilitetsfunktionen av summan.
- Bestäm den exakta probabiliteten att vinna Teddyn. Ge även svaret med flyttalsvärde.
- Bestäm den förväntade investeringen som krävs för att vinna Teddyn.
- Vad är probabiliteten att vinna en Teddy om man spelar tjugo gånger?

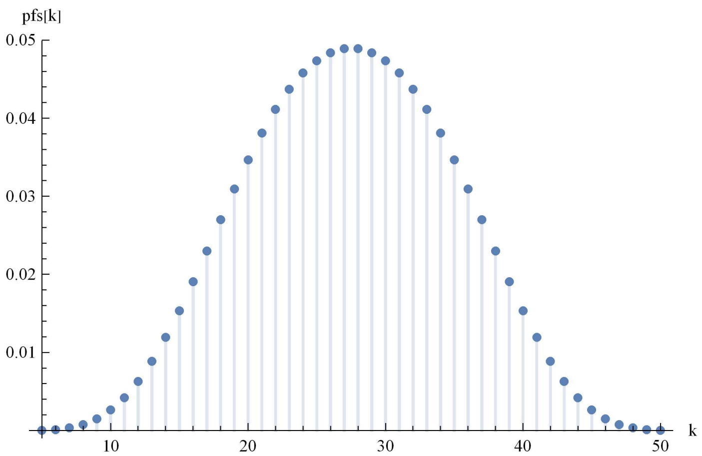
Resultat

Finn den exakta probabilitetsfunktionen av summan

För att få fram probabilitetsfunktionen krävs det att de olika utfallen för all tärningsslag summeras ihop till en enda stokastisk variabel. Summering av stokastiska variabler kallas för “Convolution”. Genom att använda den färdiga formeln DiscreteConvolve så kan vi snabbt och enklet beräkna den sammanlagda stokastiska variabeln för alla tärningar som en funktion:

$$
\begin{cases}
\frac{1}{46080} & s == 5 \,||\, s == 50 \\
\frac{1}{9216} & s == 6 \,||\, s == 49 \\
\frac{1}{3072} & s == 7 \,||\, s == 48 \\
\frac{7}{9216} & s == 8 \,||\, s == 47 \\
\frac{23}{15360} & s == 9 \,||\, s == 46 \\
\frac{121}{46080} & s == 10 \,||\, s == 45 \\
\frac{97}{23040} & s == 11 \,||\, s == 44 \\
\frac{29}{4608} & s == 12 \,||\, s == 43 \\
\frac{409}{46080} & s == 13 \,||\, s == 42 \\
\frac{61}{5120} & s == 14 \,||\, s == 41 \\
\frac{707}{46080} & s == 15 \,||\, s == 40 \\
\frac{293}{15360} & s == 16 \,||\, s == 39 \\
\frac{53}{2304} & s == 17 \,||\, s == 38 \\
\frac{311}{11520} & s == 18 \,||\, s == 37 \\
\frac{95}{3072} & s == 19 \,||\, s == 36 \\
\frac{1597}{46080} & s == 20 \,||\, s == 35 \\
\frac{39}{1024} & s == 21 \,||\, s == 34 \\
\frac{379}{9216} & s == 22 \,||\, s == 33 \\
\frac{1007}{23040} & s == 23 \,||\, s == 32 \\
\frac{211}{4608} & s == 24 \,||\, s == 31 \\
\frac{1091}{23040} & s == 25 \,||\, s == 30 \\
\frac{223}{4608} & s == 26 \,||\, s == 29 \\
\frac{1127}{23040} & 26 < s < 29 \\
0 & \text{True}
\end{cases}
$$

Funktionen kan även plottas för att ge en klarare bild av dess beteende.



Bestäm den exakta probabiliteten att vinna Teddyn. Ge även svaret med flyttalsvärde

För att vinna Teddyn så behöver man enbart beräkna den sammanlagda probabiliteten av vinstchanserna. Då vi redan beräknat den sammanlagda funktionen så blir det tämligen simpelt. Man adderar probabiliteterna mellan $5 \leq k \leq 10$ och $45 \leq k \leq 50$, vilket ger det exakta svaret $\frac{41}{3840}$. Numeriskt blir värdet 0.0106771, vilket kan tolkas som att det är strax över en procents chans att man vinner Teddyn.

Bestäm den förväntade investeringen som krävs för att vinna Teddyn.

Vi kan jämföra denna uppgift med en vanlig tärning. Alla sidor har $\frac{1}{6}$ chans att dyka upp. Därmed kan man förvänta sig att man får sin önskade sida efter sex kast. Matematiskt kan man skriva det som $1 \Big/ \frac{1}{6} = 6.$
Med samma logik kan man beräkna investeringen som krävs för att vinna Teddyn. Det finns en $\frac{41}{3840}$ chans att vinna Teddyn. Därmed kräv det $1/\frac{41}{3840}$ antal kast för att vinna Teddyn.

$$1 \Big/ \frac{41}{3840} = \frac{3840}{41} = 93.6585 \approx 94 \text{ kast}$$

Eftersom varje kast kostar 2 Euro så blir den förväntade investeringen 188 euro.

Vad är probabiliteten att vinna en Teddy om man spelar tjugo gånger?

Genom att ta den sammanlagda summan av alla utfall (100% eller 1) och subtrahera förlustchansen för 20 kast så kan vi beräkna vinstchansen för 20 kast: **`1 - ((1 - p)^20) = 0.1932...`**
Vilket kan tolkas som ~20% att vinna teddyn om man spelar 20 gånger, vilket passar in på vår tidigare beräkning i den föregående uppgiften, då ett kast har en cirka 1% chans att vinna Teddyn.

Kod

Uppgift 2: Normalfördelning

Sammanfattning

Uppgift

I denna uppgift ska vi fortsätta med uppgift 1 men vi ska använda normalfördelning för att svara på deluppgifterna nedan:

- Bestäm förväntningsvärden och standardavvikelsern för fördelningen i Teddy-fallet.
- Bestäm sannolikheten för teddypriset med normal distribution.
- Jämför och kommentera resultaten mellan första och andra uppgiften.

Resultat

Väntevärde och standardavvikelsern

Vi fick att väntevärdet blev:

$$\mu = E(X) = \frac{55}{2} = 27.5$$

och standardavvikelsen blev:

$$\sigma = \sqrt{\frac{655}{12}} = 7.38805$$

Sannolikheten för teddypriset med normal distribution.

Vi behövde bara normalfördelningens sannolikhetsfunktion $\mathrm{cdf}_x(x)$ för att få ut sannolikhet att man vinner en teddy.
Vi fick sannolikheten $0.015528 \simeq 1.55\%$ för att vinna en teddy.

Uträkning och diskussion

Väntevärdet och standardavvikelsern

Väntevärdet för en diskret stokastisk variable X defineras som:

$$\mu = E(X) = \sum_{\forall x} x \, \mathrm{pf}_X(x)$$

Där s.v. X är summan av ett kast och $\mathrm{pf}_X(x)$ är sannolikheten för att få summan X.

Med formeln fick vi väntevärdet $\mu = \frac{55}{2} = 27.5$

För att få ut standardavvikelse σ måste vi först få ut variansen σ^2:

$$\sigma^2 = E\left((X - \mu)^2\right) =$$
$$E\left(X^2 - 2\,X\mu + \mu^2\right) = E\left(X^2\right) - 2\,E(X)\,\mu + \mu^2 =$$
$$E\left(X^2\right) - \mu^2$$

$$E\left(X^2\right) = \sum_{\forall x} x^2 \, \mathrm{pf}_X(x) = \frac{4865}{6} \simeq 810.833$$

$$\mu^2 = (27.5)^2 = 756.25$$

$$\sigma^2 = \frac{4865}{6} - 756.25 = \frac{655}{12} \simeq 54.5833$$

Om man tar roten ur variansen så får man standardavvikelsen:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\left(E\left(X^2\right) - \mu^2\right)}$$

$$\sigma = \sqrt{\frac{655}{12}} = 7.38805$$

Sannolikheten för teddypriset med normal distribution

Nu kan vi få ut normalfördelningen genom att lägga in väntevärdet och standardavvikelsen i $N(\mu, \sigma)$:

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$$

$$\phi(x) = \int_{-\infty}^{x} \varphi(t)\, dt$$

$$\mathrm{pdf}_t(t) = \frac{1}{\sigma} \phi\left[\frac{t-u}{\sigma}\right] = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi}\,\sigma}$$

$$\mathrm{cdf}_x(x) = \phi\left[\frac{x-u}{\sigma}\right] = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\mu-x}{\sqrt{2}\,\sigma}\right)$$

Då får vi ut fördelningfunktionen $P(X \leq x) = \mathrm{cdf}_x(x)$ och täthetsfunktionen $P(a \leq X \leq b) = \mathrm{pdf}_t(t)$.

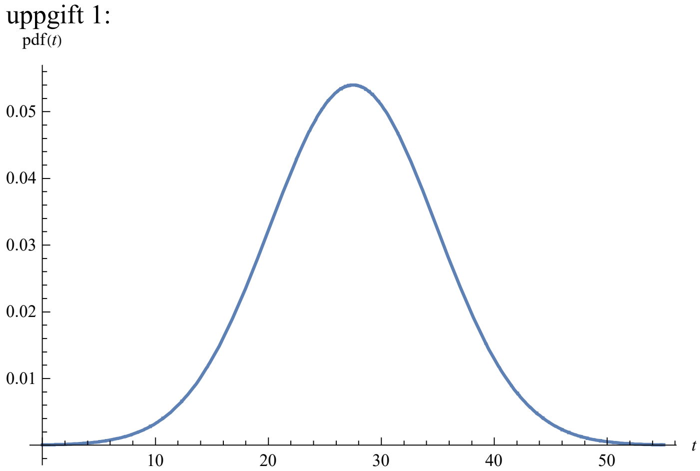
Med fördelningfunktionen $\mathrm{cdf}_x(x)$ kan vi få ut sannolikheten att man vinner en teddy:

$$P(\mathrm{Teddy}) = (\mathrm{cdf}_x(10) - \mathrm{cdf}_x(5)) + (\mathrm{cdf}_x(50) - \mathrm{cdf}_x(45)) \;.$$

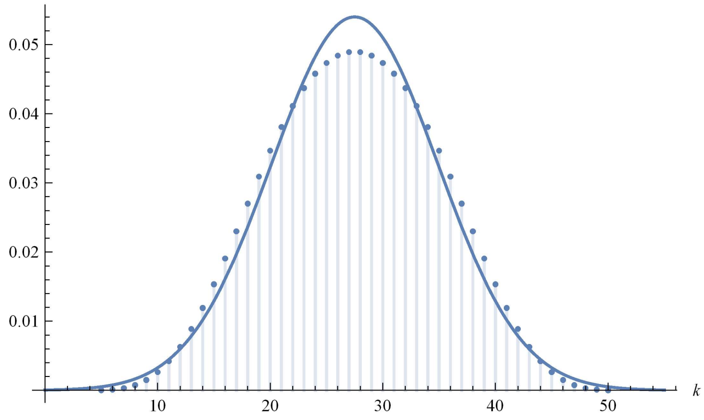
$$P(\mathrm{Teddy}) = 0.015528 \simeq 1.55\ \%.$$

Lite diskussion kring normalfördelningen

Om man ritar ut normalfördelningens täthetsfunktion så får man graf som är ganska lika diskreta fördelningen i uppgift 1:



Nedan kan vi se skillnaden grafiskt mellan normalfördelningen och diskreta fördelningen:



Man kan se att deras kurvor ganska nära varandra men normalfördelningen högsta punkt är lite högre än diskreta fördelningens punkt.

Hur fick vi en funktion som har nästan samma egenskaper som diskreta funktionen?

Tärningskast spelet som vi undersökte har observationsvärden som följer ett visst mönster. Många olika oberoende fenomen i naturen och samhället följer samma mönster, som kallas för normalfördelning.

Normalfördelningen ändras beroende på väntevärdet μ och standardavikelse σ.
Väntevärdet bestämmer positionen för grafens centrum och standardavvikelsen bestämmer grafens höjd, dvs. storleken på observationsvärdens utspriddning från väntevärdet.
Ju mer data man har desto närmare man får en funktion som har nästa samma egenskaper som normalfördelningen.

Kod