Tal: David på Folksam har försäkrat 1000 hus mot brand, sannolikheten att ett hus brinner upp är 0.0001. Vad är sannolikheten att högst 2 hus brinner upp?

L: Låt X vara antal hus som brinner upp, vad är $P(X \leq 2)$?

Förra gången sa vi att X var Bin(1000, 0.0001)-fördelad och kom fram till att

$$P(X \leq 2) \approx 0.99984578$$

Nu ska vi använda Poissonapproximationen så vi antar att $X \sim Poi(\lambda)$ där $\lambda = np = 1000 \cdot 0.0001 = 0.1$.

Vi har då att

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) =$$

$$= \frac{0.1^0}{0!} \cdot e^{-0.1} + \frac{0.1^1}{1!} e^{-0.1} + \frac{0.1^2}{2!} e^{-0.1} =$$

$$= \ldots \approx 0.99984535$$

Skillnaden mellan detta svar och det förra är ungefär 0.0000005 !

Mer om väntevärden

Kom ihåg definitionen för väntevärdet:

$$E(X) = \sum x P(X = x)$$

Följande egenskaper är användbara:

1) Om X är en slumpvariabel och a är en konstant så gäller att

$$E(aX) = aE(X)$$

2) Om X och Y är två slumpvariabler så gäller att

$$E(X+Y) = E(X) + E(Y)$$

OBS! Dessa två regler gäller för både diskreta och kontinuerliga slumpvariabler!

Om $X_1, X_2, \ldots, X_n$ har samma fördelning och

om $\mu = E(X_1) = E(X_2) = \ldots = E(X_n)$ så gäller

att

$$E\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) \overset{1)}{=} \frac{1}{n} E(X_1 + X_2 + \ldots + X_n) =$$

$$\overset{2)}{=} \frac{1}{n}\left(E(X_1) + E(X_2) + \ldots + E(X_n)\right) = \frac{1}{n}(\mu + \mu + \ldots + \mu) =$$

$$= \frac{n\mu}{n} = \mu$$

Ex: Vi slår en tärning 1000 gånger och är intresserade av medelvärdet. Låt X_k vara resultatet av k slag, så de 1000 slagen representeras av $X_1, X_2, \ldots, X_{1000}$ och medevärdet av slagen är då

$$\frac{X_1 + X_2 + \ldots + X_{1000}}{1000} = \frac{1}{1000} \sum_{k=1}^{1000} X_k .$$

Vi har även att

$$E(X_1) = E(X_2) = \ldots = E(X_{1000}) = 3.5$$

och enligt räkningen ovan har vi då att

$$E\left(\frac{X_1 + X_2 + \ldots + X_{1000}}{1000}\right) = 3.5$$

Tal: Johanna tränar frisparkar på träning. Hon skjuter 100 (oberoende) frisparkar och sannolikheten att en viss frispark går i mål är 0.1 (samma för alla 100 frisparkar). Vad är det förväntade antalet gjorda mål?

L: Låt $X_k = \begin{cases} 1 & \text{om frispark } k \text{ blir mål} \\ 0 & \text{annars} \end{cases}$

Om $X = X_1 + X_2 + \ldots + X_{100}$, vad är X?

Jo, antalet gjorda mål totalt!

Vi söker alltså $E(X)$. Vi har att

$$E(X) = E(X_1 + X_2 + \ldots + X_{100}) \overset{2)}{=} E(X_1) + E(X_2) + \ldots + E(X_{100}).$$

Vi har också att

$$E(X_k) = 0 \cdot P(X_k=0) + 1 \cdot P(X_k=1) =$$

$$= P(X_k=1) = 0.1, \quad \text{för alla } k.$$

alltså har vi att

$$E(X) = \underbrace{0.1+0.1+\ldots+0.1}_{100 \text{ ggr}} = 100 \cdot 0.1 = \underline{\underline{10}}$$

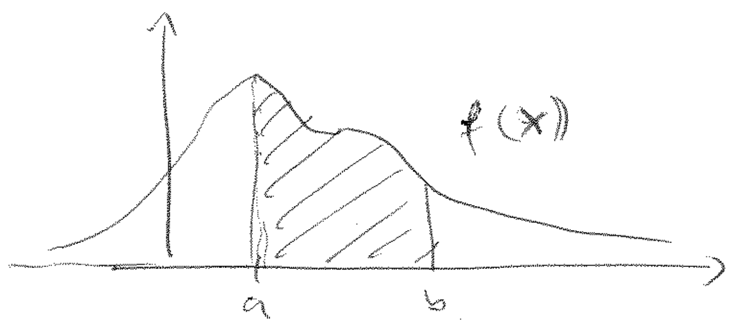
Observera att X i detta exempel är binomialfördelad med $n=100$ och $p=0.1$.

Kontinuerliga slumpvariabler

Kom ihåg: En diskret slumpvariabel har en sannolikhetsfunktion som gör att vi kan räkna ut saker som $E(X)$, $Var(X)$, $P(-5 \leq X \leq \frac{7}{2})$, osv.

Motsvarigheten för en kontinuerlig slumvariabel är dess täthetsfunktion (probability density function), som betecknas med $f(x)$ (eller $f_X(x)$).

Det kan t.ex. se ut så här



Vi har då att

$$P(a < X < b) = \text{"arean under grafen mellan a och b"}$$
$$= \int_a^b f(x)\,dx$$

<u>OBS</u>: För en kontinuerlig slumpvariabel gäller alltid att $P(X = a) = 0$ för alla a.
Arean över en enda punkt är noll!

Därför gäller det att $P(a < X < b) = P(a \leq X \leq b)$ men detta gäller generellt inte för diskreta slumpvariabler.

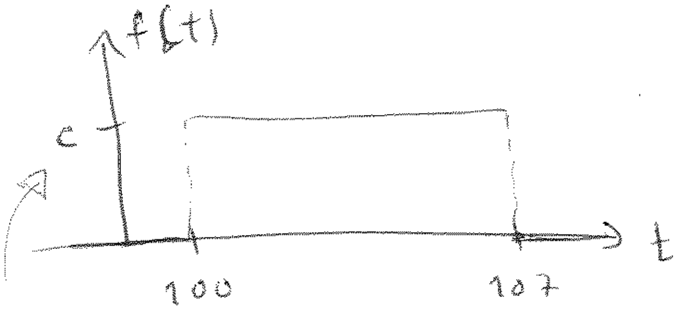
$$\boxed{\begin{array}{l}\text{En täthetsfunktion } f(x) \text{ uppfyller alltid} \\ 1)\quad f(x) \geq 0 \quad \text{för alla } x \\ 2)\quad \int_{-\infty}^{\infty} f(x)\,dx = 1\end{array}}$$

Observera att ~~integralen~~

$$\int_{-\infty}^{\infty} f(x)\,dx = P(-\infty < X < \infty)$$

Så villkor 2) kan tolkas som att "sannolikheten att X antar något värde är 1".

Tal: Ivars föreläsningar börjar klockan 10:00. Den totala föreläsningstiden T (inklusive rast) varierar något och har en täthetsfunktion, som $f(t)$ kan beskrivas med följande graf



Vad är värdet på konstanten c?

L: Om $f(t)$ ska vara en täthetsfunktion så måste vi ha att $\int_{-\infty}^{\infty} f(t)\,dt = 1$ dvs att arean under grafen $f(t)$ ska vara 1. Arean för rektangeln i bilden är $c \cdot (107-100) = c \cdot 7$, alltså måste vi ha att $c = \frac{1}{7}$.

Precis som för diskreta slumpvariabler så kan vi ta väntevärdet av en kontinuerlig slumpvariabel.

Definition Väntevärdet av en kontinuerlig slumpvariabel X med täthetsfunktion $f(x)$ ges av

$$E(X) = \int_{-\infty}^{\infty} x\, f(x)\,dx$$

Även här ~~[illegible]~~ kan vi beräkna väntevärdet av en funktion av en slumpvariabel:

Om X är en kontinuerlig slumpvariabel med täthetsfunktion $f(x)$ och g är en annan funktion så gäller att

$$\boxed{E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x)\, dx}$$

Observera att vi inte behöver någon ny definition på variansen. Även för kontinuerliga slumpvariabler har vi att

$$\mathrm{Var}(X) = E\left[(X - E(X))^2\right]$$

och räkneregeln

$$\mathrm{Var}(X) = E(X^2) - E(X)^2$$

gäller också.