STK1100 våren 2023

Kontinuerlige stokastiske variabler
Forventning og varians
Momentgenererende funksjoner

Svarer til avsnittene 4.1 og 4.2 i læreboka

Matematisk institutt
Universitetet i Oslo

En diskret stokastisk variabel kan anta endelig mange eller tellbart uendelig mange mulige verdier.

En kontinuerlig stokastisk variabel kan (i prinsippet) anta alle verdier i et intervall på tallinja (eventuelt hele tallinja).

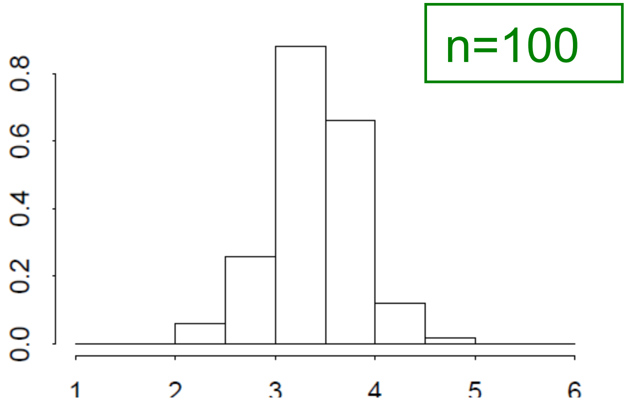
Eksempel: X = «vekt til nyfødt jente».

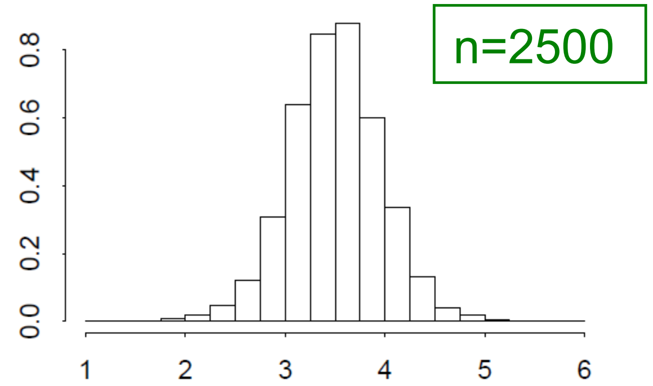
Sannsynlighetsfordelingen for en diskret stokastisk variabel er gitt ved punktsannsynligheten $p(x)=P(X=x)$
(jf. kapittel 3 i læreboka).

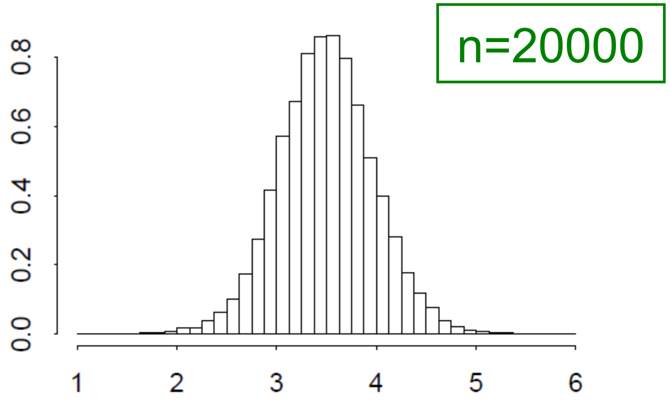
Vi kan ikke gi sannsynlighetsfordelingen for en kontinuerlig stokastisk variabel på samme måte.

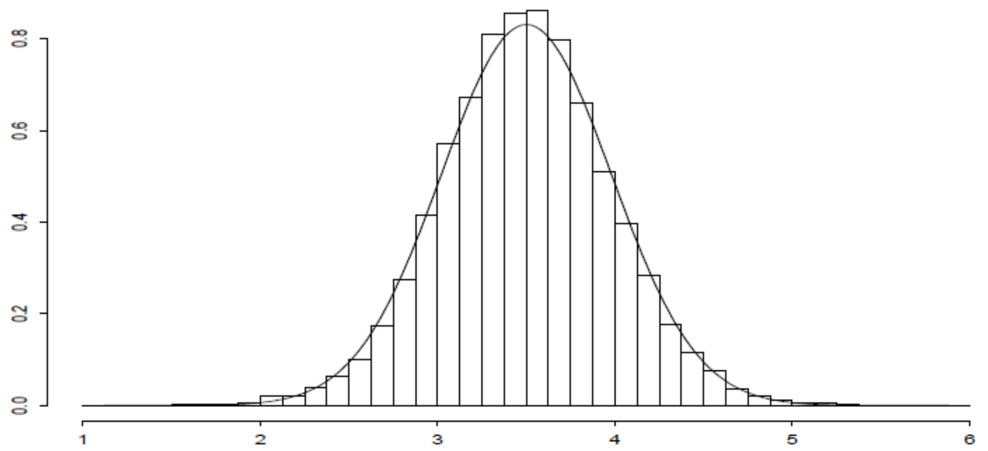
Vil se hvordan vi kan angi sannsynlighetsfordelingen til X = «vekt til nyfødt jente».

Tegner histogrammer for fødselsvektene til tilfeldige utvalg av 100, 2500 og 20000 «fullbårne» jenter (svangerskapslengde minst 37 uker). Histogrammene er normert slik at arealet av en søyle er lik den relative frekvensen for det intervallet søylen dekker.









Når *n* er stor, kan histogrammet tilnærmes med funksjonen

$$f(x) = \frac{1}{0.48\sqrt{2\pi}} e^{-\frac{1}{2\cdot 0.48^2}(x-3.5)^2}$$

$f(x)$ kalles sannsynlighetstettheten til X og svarer til histogrammet for «uendelig mange» fødselsvekter.

Vi har da at $P(a \leq X \leq b) = \int_a^b f(x)dx$

Generelt har vi følgende:

La X være en kontinuerlig stokastisk variabel med sannsynlighetstetthet $f(x)$. Da er

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Ofte sier vi tetthet i stedet for sannsynlighetstetthet

Merk at vi for enhver sannsynlighetstetthet har at

- $f(x) \geq 0$ for alle x
- $\int_{-\infty}^{\infty} f(x)dx = 1$

La X være en kontinuerlig stokastisk variabel med sannsynlighetstetthet $f(x)$. Da er

$$P(X=c) = \lim_{\varepsilon \to 0} P(c-\varepsilon \le X \le c+\varepsilon) = \lim_{\varepsilon \to 0} \int_{c-\varepsilon}^{c+\varepsilon} f(x)dx = 0$$

Når X er kontinuerlig fordelt, har vi at

$$P(a \le X \le b) = P(a < X < b) = P(a < X \le b) = P(a \le X < b)$$

Slik er det ikke hvis X er diskret fordelt

Eksempel: Uniform fordeling

Detaljer på forelesningen (jfr. læreboka side 192)

Eksempel: Flomhastighet

Detaljer på forelesningen (jfr. eksempel 4.5 i læreboka)

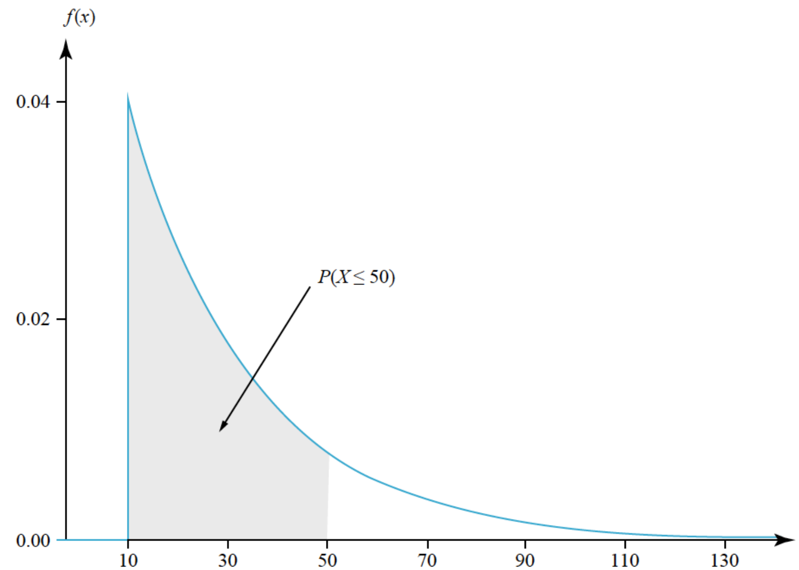


Figure 4.4 The density curve for flood rate in Example 4.5

Eksempel 4.5 i læreboka – Flomhastighet over en viss grense (10 m³/s) – artikkel i *Structural Safety* (2017):

$$f(x) = .04e^{-.04(x-10)} \quad x \geq 10$$

For en kontinuerlig stokastisk variabel X er den kumulative fordelingsfunksjonen gitt ved

$$F(x) = P(X \leq x) \quad = \int_{-\infty}^{x} f(y)\,dy$$

Omvendt har vi at $\quad F'(x) = f(x)$

Merk at: $\quad P(X > a) = 1 - F(a)$

$$P(a \leq X \leq b) = F(b) - F(a)$$

Eksempel: Uniform fordeling (forts.)
Detaljer på forelesningen

Eksempel: Flomhastighet (forts.)
Detaljer på forelesningen

La p være et tall mellom 0 og 1

100p-persentilen $\eta(p)$ i fordelingen til X er gitt ved

$$p = F\left[\eta(p)\right] = \int_{-\infty}^{\eta(p)} f(y)dy$$

Persentilene kalles også kvantiler eller fraktiler

Medianen $\tilde{\mu}$ er 50-persentilen, så $0.50 = F(\tilde{\mu})$

Eksempel: Flomhastighet (forts.)

Detaljer på forelesningen

La X være en kontinuerlig stokastisk variabel med sannsynlighetstetthet $f(x)$. Da er forventningsverdien (forventningen) til X gitt ved

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

Forventningsverdien eksisterer så sant

$$\int_{-\infty}^{\infty} |x| f(x)dx < \infty$$

Eksempel: Uniform fordeling (forts.)

Detaljer på forelesningen

Eksempel: Flomhastighet (forts.)

Detaljer på forelesningen

Regneregler for forventning

$$\mu_{h(X)} = E[h(X)] = \int_{-\infty}^{\infty} h(x) \cdot f(x) dx$$

$$E(aX + b) = a \cdot E(X) + b$$

Eksempel: "Broken stick" modellen

Detaljer på forelesningen (jfr. eksempel 4.12)

La X være en kontinuerlig stokastisk variabel med sannsynlighetstetthet $f(x)$ og forventnings-verdi μ. Da er variansen til X gitt ved

$$\sigma_X^2 = V(X) = E\left\{(X-\mu)^2\right\} \quad = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\,dx$$

Standardavviket til en stokastisk variabel X er gitt ved $\sigma_X = SD(X) = \sqrt{V(X)}$

Regneregler for varians

$$V(X) = E(X^2) - \left[E(X)\right]^2$$

$$V(aX+b) = a^2 \cdot V(X)$$

Momentgenererende funksjon (mgf)

La X være en kontinuerlig stokastisk variabel med sannsynlighetstetthet $f(x)$

Den momentgenererende funksjon for X er gitt ved

$$M_X(t) = E\left(e^{tX}\right) = \int_{-\infty}^{\infty} e^{tx} \cdot f(x)dx$$

Den momentgenererende funksjonen eksisterer hvis det fins et tall $t_0 > 0$ slik at

$$\int_{-\infty}^{\infty} e^{tx} \cdot f(x)dx < \infty \quad \text{for alle} \quad t \in (-t_0, t_0)$$

Eksempel tilsvarende 4.17 i læreboka

Detaljer på forelesningen

Momentgenererende funksjoner for kontinuerlige stokastiske variabler har tilsvarende egenskaper som for diskrete stokastiske variabler.

Spesielt: Hvis $M_X(t) = M_Y(t)$ så har X og Y samme fordeling.

Andre viktige egenskaper

1) $M_X(0) = 1$

2) $M_X^{(r)}(0) = E(X^r)$

3) Hvis $R_X(t) = \ln\{M_X(t)\}$, så er

$R'_X(0) = E(X)$ og $R''_X(0) = V(X)$

Eksempel tilsvarende 4.18 i læreboka

Detaljer på forelesningen