Kapittel 6, Kontinuerlege Sannsynsfordelingar

- Sjå på eit utval av ofte brukte kontinuerlege sannsynsfordelingar
  - Uniform **Onsdag**
  - Normal **Onsdag**
  - Eksponensial **I dag**
  - Gamma **I dag**
  - Kji-kvadrat **I dag**
  - Student-T (Kap 8.6)
- Nokre eigenskapar
  - $E(X)$ og $Var(X)$
- Samanheng mellom fordelingar **I dag**

REP Kontinuerleg sannsynsfordeling

Definisjon

Funksjonen $f(x)$ definert for alle reelle tal $x \in \Re$ blir kalla sannsynsfordelinga til den kontinuerlege stok. var. X dersom

- $0 \leq f(x)$
- $\int_{-\infty}^{\infty} f(x)dx = 1$
- $P(a \leq X \leq b) = \int_a^b f(x)dx$

REP Poissonprosess og fordeling

Poissonprosess:

- Talet på hendingar som inntreff i eit intervall er uavhengig av talet på hendingar som inntreff i disjunkte intervall.
- Sannsynet for at ei hending inntreff i eit lite intervall er lineært med lengda på intervallet, og er uavhengig av om det inntreff hendingar før eller etter intervallet.
- Sannsynet for at meir enn ei hending inntreff i eit lite intervall er neglisjerbart.

Poissonfordeling

$$p(x; \lambda t) = \frac{\exp(-\lambda t)(\lambda t)^x}{x!}$$

REP: Kummulativ sannsynsfordeling

Definisjon 3.5

Den *kummulative fordelinlgsfunksjonen* $F(x)$ *for ein stok. var.* X *er:*

$$F(x) = P(X \leq x)$$

- *Diskret:* $F(x) = \sum_{t \leq x} f(t)$
- *Kontinuerleg* $F(x) = \int_{-\infty}^{x} f(t) dt$

Eksponensialfordeling

Den stokastiske variabelen X er eksponensial fordelt dersom

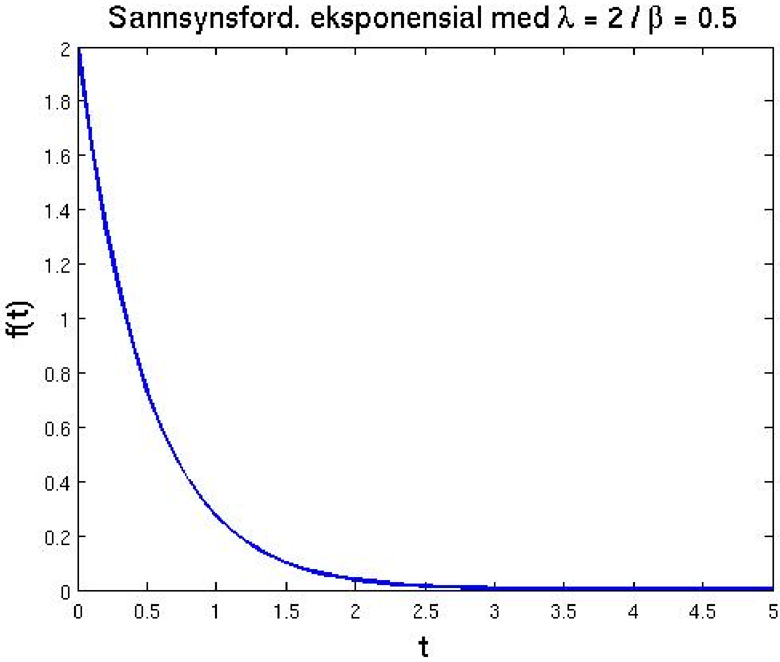
$$f(x;\beta) = \frac{1}{\beta}\exp(-x/\beta)$$

for $x > 0$ og 0 ellers.
Kummulativ fordeling:

$$F(x;\beta) = P(X < x) = 1 - \exp(-x/\beta)$$

for $x > 0$.

$E(X) = \beta$
$Var(X) = \beta^2$



Sannsynsford. eksponensial med λ = 2 / β = 0.5
f(t)
t
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

Du er vedlikehaldansvarleg for ein installasjon. Du veit at levetida til ei viktig pumpe er eksponensialfordelt med forventa levetid 10 år. Det er gått 9 år, og sjefen spør om ikkje denne pumpa bør skiftast ut. Kva svarer du?

Er levetid for menneske eksponensialfordelt?

Tabell frå SSB

Gammafordeling

Den stokastiske variabelen X er gamma fordelt dersom

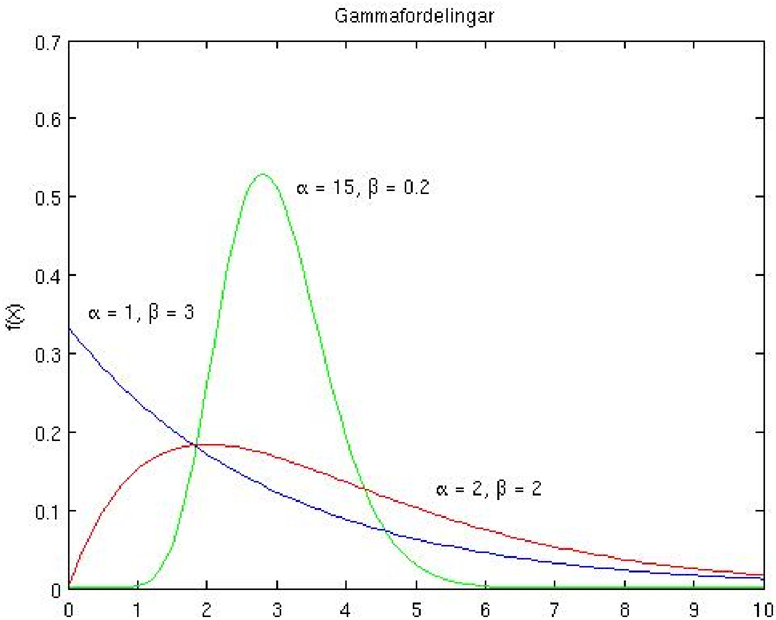
$$f(x;\beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} \exp(-x/\beta)$$

for $x > 0$ og 0 ellers der $\Gamma()$ er gamma-funksjonen

$E(X) = \alpha\beta$
$Var(X) = \alpha\beta^2$

- α: *shape*
- β: *scale*

Gamma-funksjonen: $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} \exp(-x) dx$. For α heiltal: $\Gamma(\alpha) = (\alpha-1)!$
PS: Og vanleg å bruke parametrisering der $\beta \to 1/\beta$!

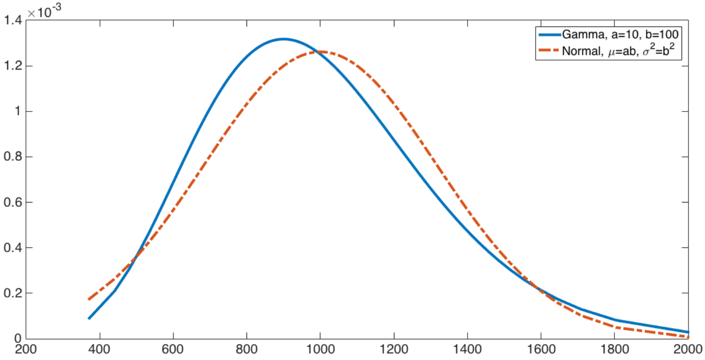

Gammafordelingar
f(x)
α = 15, β = 0.2
α = 1, β = 3
α = 2, β = 2
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0 1 2 3 4 5 6 7 8 9 10

Spesialtilfeller gammafordeling

$$X \sim Ga(\alpha, \beta)$$

$$\begin{aligned}
\alpha = 1: &\ \text{Eksponensial med parameter } \beta \\
\beta = 2 \text{ og } \alpha = \nu/2: &\ \chi^2\text{-fordelt med parameter } \nu \\
\alpha \to \infty: &\ \Rightarrow X \to N(\alpha\beta, \alpha\beta^2)
\end{aligned}$$

Gamma mot Normal



Bruksområde gammafordeling

- Når andre ikkje passar (f.eks. mengde nedbør per månad)
- Kan vise at ventetida til hending nr α i ein Poisson-prosess med intensitet λ er $Ga(\alpha, \beta = 1/\lambda)$
- χ^2-fordeling: I samband med estimering av varians.

Kap 6.7: χ^2 fordeling

Notasjon: $X \sim \chi^2_\nu$

ν: Parameteren i χ^2-fordelinga. Blir kalla *talet på fridomsgrader*

Eigenskapar:

- $E(X) = \nu$
- $Var(X) = 2\nu$
- Dersom $Z \sim N(0, 1)$, så er $Z^2 \sim \chi^2_\nu$
- Brukt som utvalsfordeling til estimert varians (kap 8.5)

REP: 5.2 Bernoulli prosess og binomisk fordeling

Bernoulli prosess

1. n uavhengige forsøk
2. Kvart forsøk resulterer i suksess, $I_i = 1$ eller ikkje-suksess $I_i = 0$.
3. Suksess-sannsynet $p = P(I_i = 1)$ er konstant.

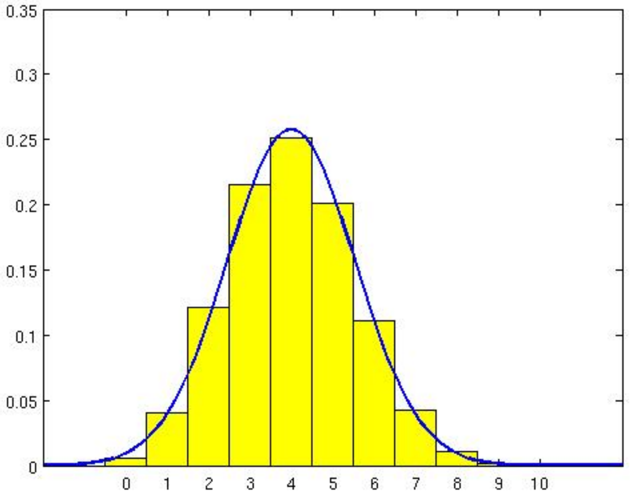
Bernoulli prosess = trekking med tilbakelegging

Binomisk fordeling

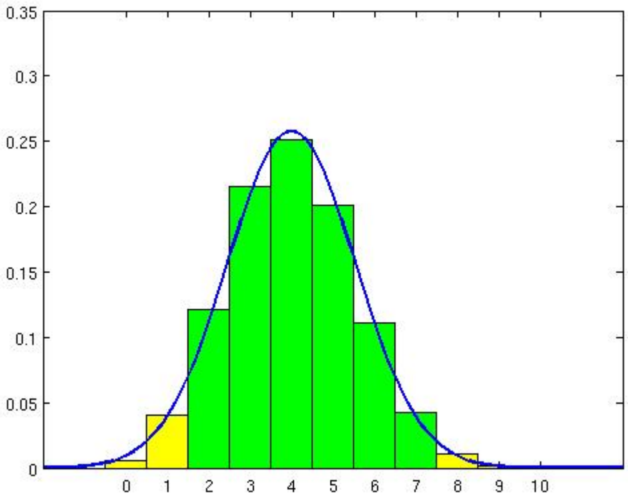
Ser på antall suksess i ein Bernoulli prosess, $X = \sum_{i=1}^{n} I_i$. X er då binomisk fordelt;

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

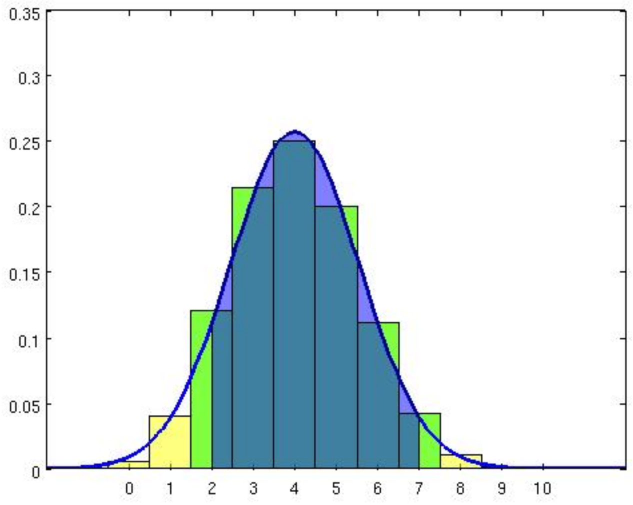
Binomisk: $p = 0.4$ og $n = 10$ Normal: forventning og var. som i binomisk.



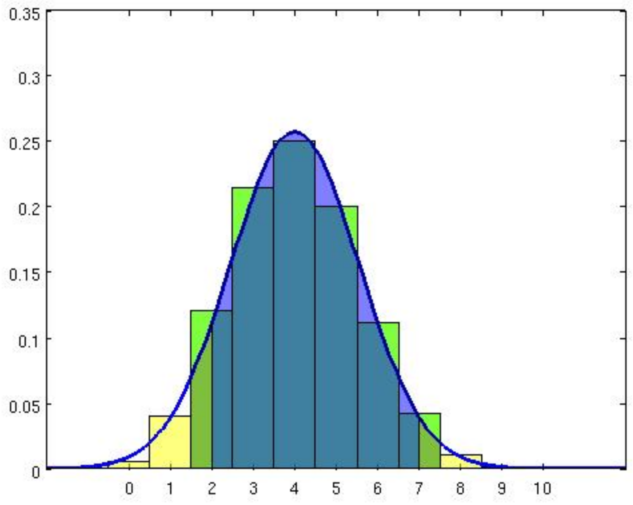
$P(2 \leq X \leq 7)$



Normaltilnærming $P(2 \leq X \leq 7)$



Normaltilnærming med halvkorreksjon



- Urne med N kuler.
- k blåe kuler (suksess)
- $N-k$ raude kuler (ikkje-suksess)
- Trekker n kuler
- X er antall av dei trekte som er blåe (suksess).

Definisjon hyper-geometrisk

$$f(x; N, n, k) = h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-k}}{\binom{N}{n}}$$

1. Talet på hendingar som inntreff i eit tidsintervall er uavhengig av talet på hendingar som inntreff i disjunkte tidsintervall.
2. Sannsynet for at ei hending inntreff i eit kort tidsintervall er lineært med lengda på tidsintervallet, og er uavhengig av om det inntreff hendingar før eller etter intervallet.
3. Sannsynet for at det inntreff meir enn ei hending innanfor eit lite tidsintervall er neglisjerbart.

PS: Kan bytte ut tid med distanse, areal, volum.

Poissonfordeling

La X vere talet på hendingar i eit tidsintervall t. X er poissonfordelt dersom sannsynsfordelinga er

$$f(x; \lambda t) = p(x; \lambda t) = \frac{\exp(-\lambda t)(\lambda t)^x}{x!}$$

der λ er gjennomsnittleg antall hendingar per tidseining (f.eks. time).

Oppsummering

Normalfordeling

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi}\sigma}\exp(-\frac{1}{2\sigma^2}(x-\mu)^2)$$

for $-\infty < x < \infty$

- $E(X) = \mu$ og $Var(X) = \sigma^2$
- $Y = a + bX$, $Y \sim N(a + bE(X), b^2 Var(X))$
- Z er standard normalfordelt dersom $Z \sim N(0, 1)$
- Kummulativ fordeling for standard normalfordeling er tabulert.

- Sjå på eit utval av ofte brukte kontinuerlege sannsynsfordelingar
  - Uniform
  - Normal
  - Eksponensial
  - Gamma
  - (Kji-kvadrat)
  - (Student-T)
- Samanheng mellom fordelingar
- Nokre eigenskapar
  - $E(X)$ og $Var(X)$