- Sjå på eit utval av ofte brukte diskret sannsynsfordelingar
  - Binomisk
  - Multinomisk
  - Hypergeometrisk
  - Geometrisk
  - Negativ binomisk
  - Poisson
- Prosserar desse beskriv
  - Bernoulli prosess
  - Poisson prosess
- Nokre eigenskapar
  - $E(X)$ og $Var(X)$

Diskret stokastisk variabel

Eit stokastisk variabel som gjev eit endeleg eller tellbart antall utfall.

Diskret

Paret $(x, f(x))$ blir kalla sannsynsfordelinga til den diskret stok. var. X dersom

- $0 \leq f(x)$
- $\sum_{\forall x} f(x) = 1$ (summen over alle mogelege x)
- $f(x) = P(X = x)$

EKSEMPEL Terningkast

x	1	2	3	4	5	6
$f(x)$	1/6	1/6	1/6	1/6	1/6	1/6

REP: Binominal-koeffisienten (teo 2.6)

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

Antall moglege kombinasjonar dersom ein trekker x individ frå ein popuasjon på n, *utan tilbakelegging*, og når ein ikkje bryr seg om ordninga (rekkefølgje dei er trekt i), dvs *uordna*.

Bernoulli prosess (= trekking med tilbakelegging)

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.

Binomisk fordeling

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

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

For plotting: Anbefaler *disttool()* i Matlab, evt denne

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* $F(x) = \int_{-\infty}^{x} f(t) dt$

Definisjon 4.1

La X vere ein stok.var. med sannsynsfordeling $f(x)$. Forventningsverdien til X er då

$$\mu = E(X) = \sum_{\forall x} x f(x)$$

dersom X er diskret.

Definisjon 4.3

La X vere ein stok. var. med forventning μ. Variansen til X er då:
For diskret X:

$$Var(X) = E((X-\mu)^2) = \sum_{\forall x} (x-\mu)^2 f(x)$$

Teorem 4.2

$$Var(X) = E(X^2) - (E(X))^2$$

REP: Forventning og varians av lineærkombinasjon

Teorem 4.5, 4.9 og 4.11

Dersom $a_0, a_1, \ldots, a_n$ er konstantar og $X_1, X_2, \ldots X_n$ stok. var., så er

$$E(a_0 + \sum_{i=1}^{n} a_i X_i) = a_0 + \sum_{i=1}^{n} a_i E(X_i)$$

og

$$Var(a_0 + \sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 Var(X_i) + 2 \sum_{i=1}^{n} \sum_{j=1}^{i-1} a_i a_j Cov(X_i, X_j)$$

for ***uavhengige*** $X_1, \ldots, X_n$ er

$$Var(a_0 + \sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2 Var(X_i)$$

Multinomisk prosess og multinomisk fordeling

Multinomisk prosess

1. n uavhengige forsøk
2. Kvart forsøk resulterer i ein av R kategoriar
3. Sannsynet for kategori r er lik i kvart forsøk (i)
   $p_i = P(\text{kategori r}) = p_r$.

Multinomisk fordeling

Ser på talet på utfall i kategori r i ein Multinomisk prosess.
$X = (X_1, X_2, \dots X_R)$ er då multinomisk fordelt;

$$f(x_1, x_2, \dots, x_R; n, p_1, p_2, \dots, p_R) = \binom{n}{x_1, x_2, \dots x_R} p_1^{x_1} p_2^{x_2} \dots p_R^{x_R}$$

Binominal-koeffisienten (teo. 2.6)

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

Antall moglege kombinasjonar dersom ein trekker x individ frå ein popuasjon på n, *utan tilbakelegging*, og når ein ikkje bryr seg om ordninga (rekkefølgje dei er trekt i), dvs *uordna*.

Multinomial koeffisient (teo 2.5)

Tilsvarende for r ulike grupper ('av suksess').

$$\binom{n!}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1! n_2! \ldots n_r!}$$

der $n_1 + n_2 + \cdots + n_r = n$.

Hypergeometrisk fordeling

- 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

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

Spørjeundersøking i studieprogram

Studieprogram har $N = 50$ studentar.
For eller i mot obligatorisk øving?
Nok å sprørje 10?

Studieprogram har $N = 50$ studentar.
For eller i mot obligatorisk øving?
Nok å sprørje 10?
Kva om $k = 20$ er for og $N - k = 30$ er i mot?
La X vere antal av dei 10 som er for. (I Matlab, M=50, K=20, N=10)
Kva er $P(X \geq 5) =$?

Forventning og varians hypergeometrisk fordeling

Teorem 5.3

Dersom X er hyper-geometrisk fordelt: h(x;N,n,k), så er forventningsverdien

$$E(X) = \mu = \frac{nk}{N}$$

og variansen

$$Var(X) = \sigma^2 = \frac{N-n}{N-1} \cdot n \cdot \frac{k}{N} \cdot (1 - \frac{k}{N})$$