

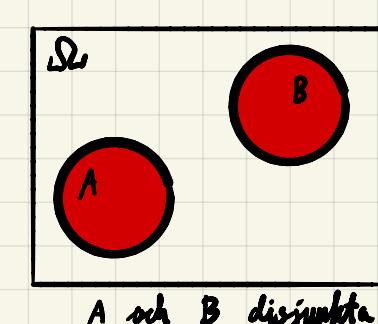
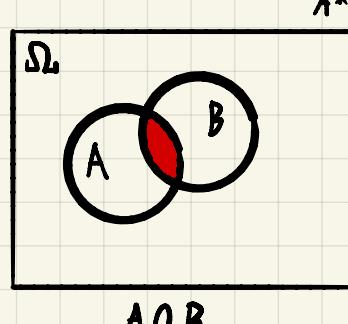
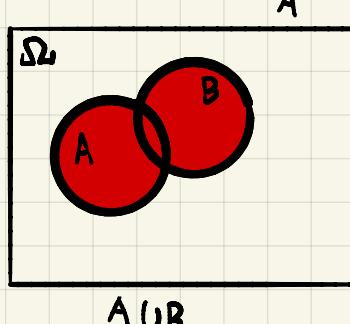
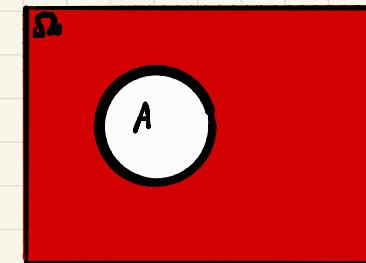
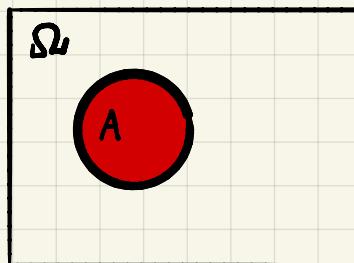
Kombinatorik

	Med återläggning	Utan återläggning	
Med ordning	n^k	$n(n-1) \cdots (n-k+1)$	$n \rightarrow$ Antal element i en mängd
utan ordning	$\binom{n+k-1}{k}$	$\binom{n}{k}$	$k \rightarrow$ Antal element som dras ur en mängd.

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \quad (\text{OBS! Ej permutation; kombination!})$$

Betingad Sannolikhet (Sannolikheten att B händer
givet att A händer)

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$



Bayes Sats

Låt A_1, \dots, A_n vara n disjunkta händelser...

$$P(A_i | B) = \frac{P(B | A_i) P(A_i)}{\sum_{j=1}^n P(B | A_j) P(A_j)}$$

$$P(A | B) = \frac{P(B | A) P(A)}{P(B)} \quad \text{← (Om } n=1, \text{ specialfall)}$$

Komplement

$$P(A) = 1 - P(A^c) \quad \boxed{A \quad A^c}$$

$$P(A \cup B) = P(A) + P(A^c \cap B)$$

$$P(A^c \cup B) = 1 - P(A \cap B^c)$$

Disjunkta

$$P(A \cup B) = P(A) + P(B)$$

$$A \cap B = \emptyset$$

Additionssatsen

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Oberoende

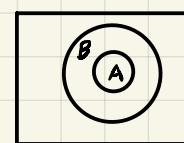
$$\begin{aligned} P(A \cap B) &= P(A) \cdot P(B) && \text{(Om } A \text{ och } B \text{ är oberoende)} \\ P(B | A) &= P(B) \end{aligned}$$

Om A och B är oberoende så är även A^c och B det.

Lagen om total sannolikhet (slumpmässigt vald...)

$$P(B) = \sum_{i=1}^n P(B | A_i) P(A_i)$$

$$\begin{aligned} A \subset B \\ \text{(zillhör)} \end{aligned}$$



Stokastiska variabler

Poissonfördelning

- Diskret fördelning som ofta används när något utspelar sig över en tidsperiod
- Sannolikhetsfunktionen för en stokastisk variabel $X \in Po(\mu)$ ges av:

$$P_x(k) = \frac{\mu^k}{k!} e^{-\mu}$$

där $k = 0, 1, 2, \dots$ och $\mu > 0$

Väntevärde: $E(X) = \mu$

Varians: $V(X) = \mu$

Täthetsfunktion $f_X(x)$

- Om X är en kontinuerlig stokastisk variabel gäller:

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

- Allmänt gäller:

$$\int_{-\infty}^{\infty} f_X(x) dx = 1$$

(då totala sannolikheten alltid är 1)

$$0 \leq f_X(x) \leq 1$$

- För kontinuerliga stokastiska variabler gäller även:

$$f(a) \neq P(X=a)$$

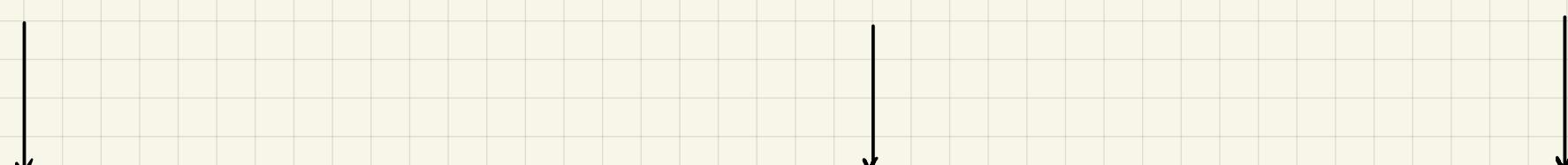
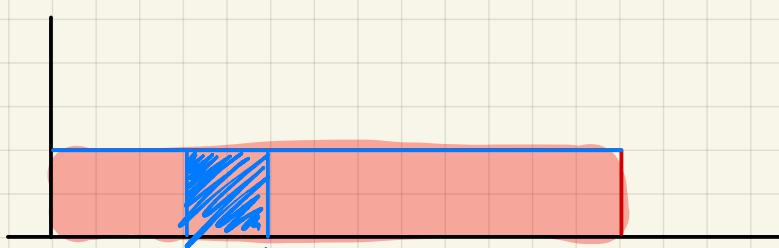
$$P(a \leq X \leq b) = P(a < X < b)$$

- Likformig fördelning

$$\int_a^b f_X(x) dx \text{ där...}$$

$$f_X(x) = \frac{1}{(b-a)}$$

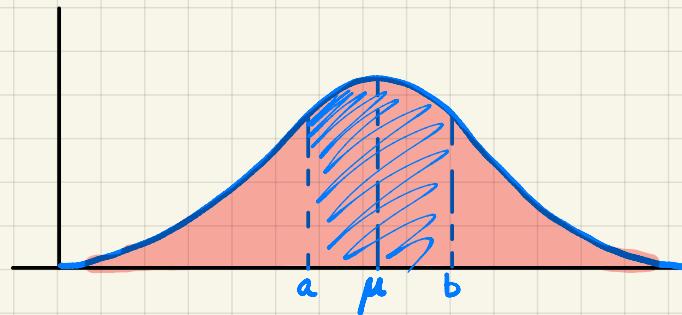
$$(a < x < b)$$



• Normalfördelning

$$\int_a^b f_x(x) dx \text{ där...}$$

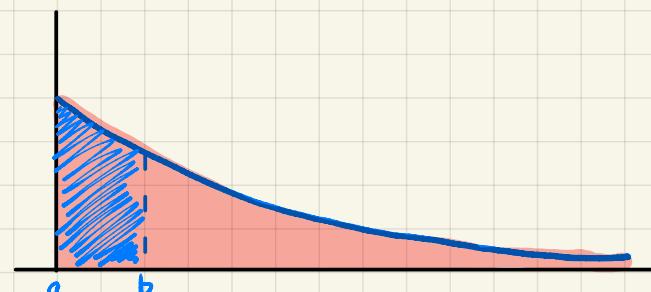
$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\mu-x)^2}{2\sigma^2}}$$



• Exponentialfördelning

$$\int_a^b f_x(x) dx \text{ där...}$$

$$f_x(x) = \lambda e^{-\lambda x}$$



$$\cdot f_x(x) = \begin{cases} \lambda e^{-\lambda x} & \text{för } x \geq 0 \\ 0 & \text{för } x < 0 \end{cases}$$

$$\cdot \text{Väntevärde: } E(X) = \frac{1}{\lambda}$$

$$\cdot \text{Varians: } V(X) = \frac{1}{\lambda^2}$$

$$\cdot P(X > A | X > B) = P(X > A - B)$$

Bionomialfördelning

• Vi vet hur många försök n vi ska göra

• Försöken är oberoende av varandra

• Det finns endast två möjliga utfall; p och $p-1$

• Samolikhefsfunktion:

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \text{ där } n \text{ är antalet försök och } 0 \leq p \leq 1$$

• Väntevärde:

$$E(X) = np$$

• Varians:

$$V(X) = np(1-p)$$

Multinomialfördelning

• Vi vet hur många försök n vi ska göra

• Försöken är oberoende av varandra

• Det finns mer än två möjliga utfall; p och $p-1$

• Samolikhefsfunktion:

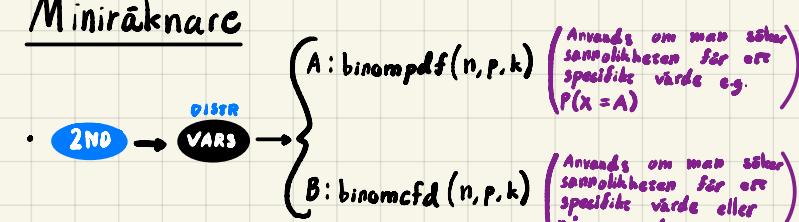
$$P_{X_1, \dots, X_r}(k_1, \dots, k_r) = \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r} \quad (\text{Varje försök kan resultera i } r \text{ möjliga utfall})$$

$$\cdot \sum_{i=1}^r p_i = 1 \quad (\text{Summan av sannolikheten av alla utfall är 1})$$



Miniräknare

• 2ND → **DISTR**



För $P(X > A)$ kan man nyttja $P(X > A) = 1 - P(X \leq A)$



$$\cdot \sum_{i=1}^r k_i = n \quad (\text{Summan av samtliga utfall är } n)$$

Fördelningsfunktion $F_X(x)$

- Fördelningsfunktionen för en kontinuerlig stokastisk variabel ges av:

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

Varians

- $V(X) = E((X - \mu)^2)$
- $V(X) = E(X^2) - (E(X))^2$

Standardavvikelse

- $\sigma = \sqrt{V(X)}$
- $V(X) = \sigma^2$

Sannolikhetsfunktion

- Om X är en **diskret** stokastisk variabel gäller:

$$P(a \leq X \leq b) = \sum_{k=a}^b P_X(k) \quad (\text{där } P_X(k) \geq 0 \text{ för alla } k \text{ och } \sum_{k=0}^{\infty} P_X(k) = 1)$$

ML-Metoden

- Diskret:

$$L(\theta) = P(X_1 = x_1, \dots, X_n = x_n; \theta) = \prod_{i=1}^n P_{X_i}(x_i; \theta)$$

- Kontinuerlig:

$$L(\theta) = f_{X_1, \dots, X_n}(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$$

- Populärt knep: maximera $\ln(L(\theta))$ i stället för $L(\theta)$.

Varians / Kovarians

- $V(aX + bY) = a^2 V(X) + b^2 V(Y) + 2ab C(X, Y)$ (**beroende** X, Y) (finns ej på
formelblad...)
- $V(aX + bY) = a^2 V(X) + b^2 V(Y)$ (**oberoende**)
- $V(aX + b) = V(aX)$

- $V(aX) = a^2 V(X)$
- $D(X) = \sqrt{V(X)}$
- $V(X) = (D(X))^2$
- $C(X, Y) = E(XY) - E(X)E(Y)$

Väntevärdesriktighet

- En punktskattning θ_{obs}^* sätgs vara väntevärdesriktig om: $E(\theta^*) = \theta$
- Vid jämförelse av skattningar används variansen av skattningarna. Skattningen med minst varians klassas som mest effektiv.
- Medelfeler är den skattade standardavvikelsen av skattningen.

Intensitet

- $\frac{\mu}{\tau \text{id}} = \text{intensitet}$

Normalfördelning

- Sats 1
Om $X \in N(\mu_x, \sigma_x)$ och $Y \in N(\mu_y, \sigma_y)$ och X och Y är oberoende gäller:

$$aX + bY \in N(a\mu_x + b\mu_y, \sqrt{(a\sigma_x)^2 + (b\sigma_y)^2})$$

$$X - Y \in N(\underbrace{\mu_x - \mu_y}_{E(X-Y)}, \underbrace{\sqrt{\sigma_x^2 + \sigma_y^2}}_{D(X-Y)})$$

- Sats 2

$$\text{Om } X \in N(\mu, \sigma)$$

$$Y = aX + b \in N(a\mu + b, |a|\sigma)$$

$$(Y = 2X \Rightarrow Y \in N(2\mu, 2\sigma))$$

- Centrala gränsvärdesatsen (CGS):

Om n är stort och X_i :na har samma fördelning och är oberoende s.v. med $E(X_i) = \mu$ och $D(X_i) = \sigma^2 > 0$ ger..

$$\sum_{i=1}^n X_i \approx N(n\mu, \sigma\sqrt{n})$$

Linjära kombinationer av P_0 , Bin

• Sats 1

Om vi har 2 oberoende s.v. X och Y som är respektive

$$X \in P_0(M_1) \quad Y \in P_0(M_2)$$

($E(X)$) ($E(Y)$)

...kan vi säga att $Z = X + Y \in P_0(M_1 + M_2)$

• Sats 2

Låt X och Y vara oberoende s.v. med samma p , sådana att:

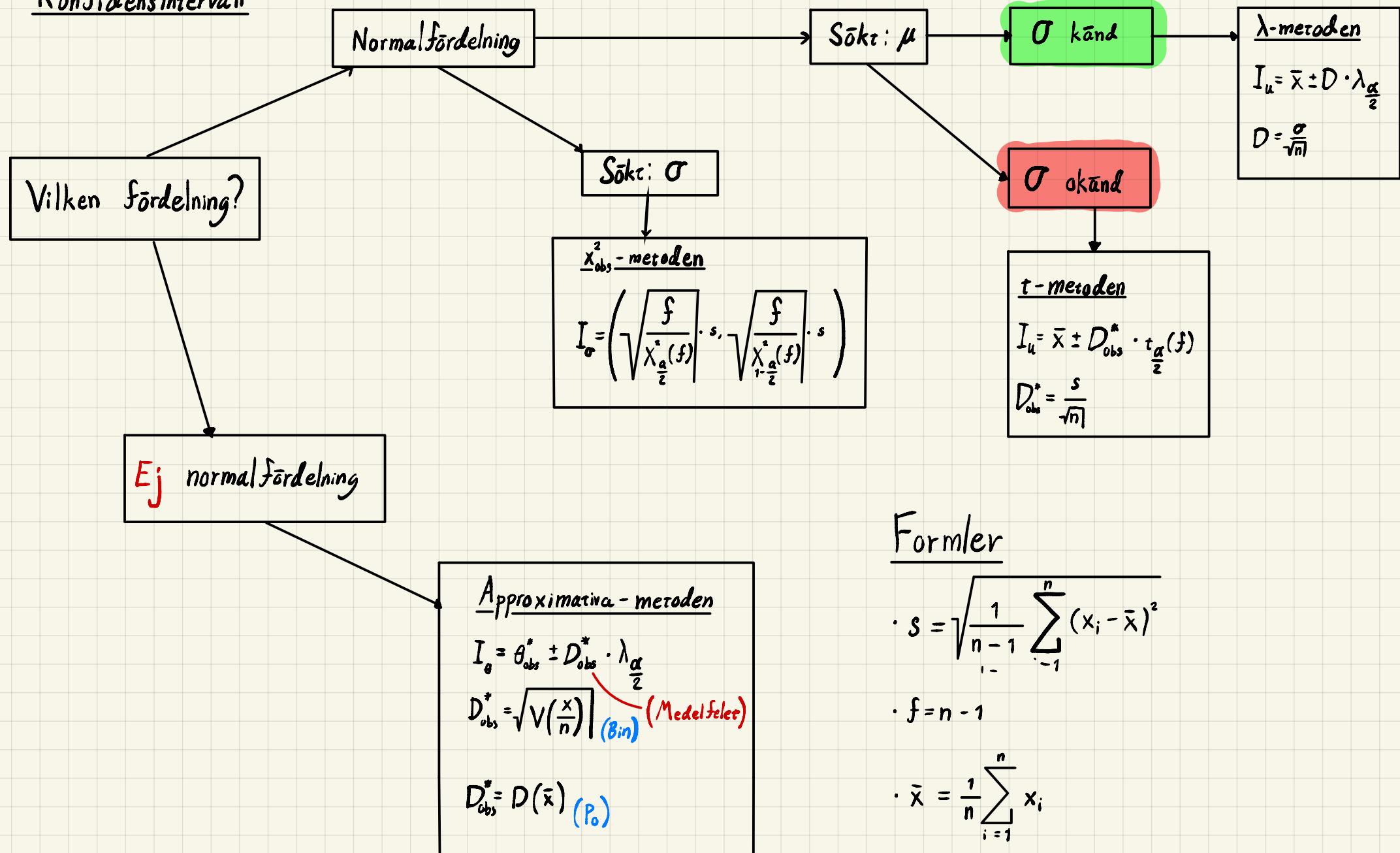
$$X \in \text{Bin}(n_x, p) \quad \text{och} \quad Y \in \text{Bin}(n_y, p)$$

...då gäller det att:

$$Z = X + Y \in \text{Bin}(n_x + n_y, p)$$

OBS! Gäller endast om X och Y har samma sannolikheter p

Konfidensintervall



Formler

$$\cdot s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\cdot f = n - 1$$

$$\cdot \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Hypotesprövning

- Den hypotes som man vill testa om man ska förkasta eller inte kallas nollhypotesen, H_0 , ska alltid kunna förkastas.
- Mothypotesen till H_0 kallas H_1 .
- Risknivå = signifikansnivå = α , $P(\text{förfästa } H_0)$ om H_0 är sann < α
- p -värde: sannolikheten att förfästa H_0 under förutsättningen att H_0 är sann givet att vi fått en viss värde på test-variabeln = observationen
 p -värde < $\alpha \Rightarrow$ förfästa H_0 , annars inte.
(signifikansnivån)
- Styrkefunktionen: $b(\theta) = P(\text{förfästa } H_0)$ om θ är rätt parametervärdet.
- Styrka hos testet för $\theta = \theta_0$: $P(\text{förfästa } H_0 : \theta = \theta_0)$ om $H_1 : \theta = \theta_0$ är sann.
- Konfidensintervall metoden
 - Om $H_0 : \mu = a$
 $H_1 : \mu \neq a$
 - Bildar ett konfidensintervall för μ och om a ligger i konfidensintervallerna så kan vi inte förfästa H_0 på den risknivån
 - Vid ensidigt test: $H_0 : \mu = a$
 $H_1 : \mu < a$ begränsar uppåt.

Medelfel

$$D(\theta) = \frac{\sigma}{\sqrt{n}}$$

$\max(X, Y)$

- $P(\max(X, Y) > a) = 1 - P(X < a)P(Y < a)$
- $P(\max(X, Y) < a) = P(X < a)P(Y < a)$

$\min(X, Y)$

- $P(\min(X, Y) < a) = 1 - P(X \geq a)P(Y \geq a)$
- $P(\min(X, Y) > a) = P(X \geq a)P(Y \geq a)$

Centrala gränsvärdessatsen

- Låt X_1, X_2, \dots, X_n vara oberoende och likafördelade stokastiska variabler från godtycklig fördelning.
- Låt Y vara summan av dessa:

$$Y_n = \sum_{i=1}^n X_i$$

Då gäller att Y_n är approximativt normalfördelad enligt:

$$Y_n \in N(n\mu, \sigma\sqrt{n})$$

Absolutbelopp

- $P(|Z| \leq a) = P(-a \leq Z \leq a)$

Additionsprincipen

- $\text{Bin}(n, p)$

$$X + Y \in \text{Bin}(\mu_x + \mu_y, p)$$

- $P_0(\mu)$

$$X + Y \in P_0(\mu_x + \mu_y)$$

- $N(\mu, \sigma)$

$$aX + bY \in N(a\mu_x + b\mu_y, \sqrt{(a\sigma_x)^2 + (b\sigma_y)^2})$$

Stickprov

- Stickprovsvarians = s^2 där $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$

χ^2 -test

- Test av given fördelning

Vi utför n stycken försök där var och ett kan urfalla på olika sätt:

A_1, A_2, \dots, A_r med sannolikheterna $P(A_1), P(A_2), \dots, P(A_r)$. Vi vill testa nollhypotesen som säger att stickprovet kommer från en given fördelning:

$$H_0: P(A_1) = p_1, P(A_2) = p_2, \dots, P(A_r) = p_r$$

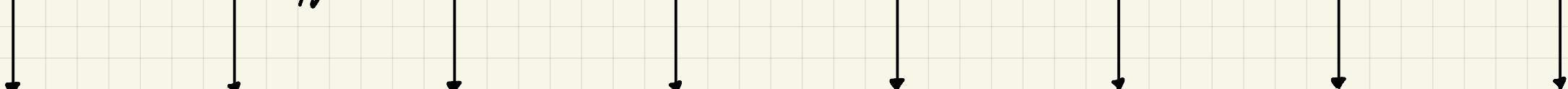
$$\sum_{j=1}^n P(A_j) = 1 \quad \sum_{j=1}^n p_j = 1 \quad \sum_{j=1}^n x_j = 1$$

där x_1, x_2, \dots, x_r är de absoluta frekvenserna för de olika urfallen i vårt försök.

Mothypotesen är att nollhypotesen inte stämmer, alltså att man kan säga att stickprovet **inte** kommer från den givna fördelningen.

En krav för att kunna använda detta test är att följande är uppfyllt för alla j :

$$np_j \geq 5$$



Vi beräknar testvariabeln Q :

$$Q = \sum_{i=1}^r \frac{(x_i - np_i)^2}{np_i}$$

H_0 förkastas om $Q > \chi^2_{\alpha}(r-1)$, då värder på χ^2 fås ur tabell.

• Homogenitetstest

Används när man utför s scycken försöksserier och vill undersöka om de kan anses homogena.

Krav för homogenitetstest är att:

$$\frac{n_i m_j}{N} \geq 5$$

Testvariabeln beräknas enligt:

$$Q = \sum_{i=1}^s \sum_{j=1}^r \frac{(x_{ij} - \frac{n_i m_j}{N})^2}{(\frac{n_i m_j}{N})}$$

H_0 förkastas om $Q > \chi^2((r-1)(s-1))$, då värder på χ^2 fås ur tabell.

• Oberoendetest

Används om man vill undersöka vädemängderna från två olika s.v. kan anses vara oberoende. Säg att den stokastiska variabeln X kan utfalla på r olika sätt A_1, A_2, \dots, A_r och den stokastiska variabeln Y kan utfalla på s olika sätt B_1, B_2, \dots, B_s .

H_0 säger att de stokastiska variablerna är oberoende.

Testvariabeln beräknas enligt:

$$Q = \sum_{i=1}^s \sum_{j=1}^r \frac{(x_{ij} - \frac{n_i m_j}{N})^2}{(\frac{n_i m_j}{N})}$$

H_0 förkastas om $Q > \chi^2((r-1)(s-1))$, då värder på χ^2 fås ur tabell.

• Test av fördelning med skattad parameter

H_0 förkastas om $Q > \chi^2(r - \underbrace{1-1}_{(2 \text{ frihetsgrader})})$, då värde på χ^2 fås ur tabell.

2 Stickprov med samma varians:

$$I_{\mu} = \bar{x} - \bar{y} \pm s \sqrt{\frac{1}{n_x} + \frac{1}{n_y}} \cdot t_{\frac{\alpha}{2}} (n_x + n_y - 2) \quad \text{där} \quad s = \sqrt{\frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2}}$$