Sannolikhetsteori och Statistik

Def (utfall): Resultatet av ett slumpmässigt försök kallas ett utfall. Betecknas $\omega_1, \omega_2, \ldots$

Def (utfallsrummet): Mängden av alla möjliga utfall kallas utfallsrummet. Betecknas Ω.

Def (händelse): En händelse är en samling av utfall. Betecknas $A, B, C, \ldots$

Def (diskret, ändligt, kontinuerligt): Om antalet utfall är ändligt eller uppräkneligt oändligt, säges Ω vara ett diskret utfallsrum. Om antalet är ändligt, säges Ω speciellt vara ett ändligt utfallsrum. Om antalet utfall icke är ändligt eller uppräkneligt oändligt, säges Ω vara ett kontinuerligt utfallsrum.

Notation: $\cup$ och $\cap$

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \ldots \cup A_n = \text{minst en av händelserna } A_1, \ldots, A_n \text{ kommer inträffa}$$

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \ldots \cap A_n = \text{alla händelserna } A_1, \ldots, A_n \text{ inträffar.}$$

Def (parvis oförenliga): Händelserna $A_1, \ldots, A_n$ säges vara parvis oförenliga om alla par A_i och A_j är oförenliga (disjunkta), dvs om det är omöjligt att två eller flera av händelserna inträffar samtidigt.

Def (komplement): Komplementet till A är $A^* = \Omega \setminus A$. Detta kan också betecknas som $\complement A$.

SATS 2.α (De Morgans lagar): Låt A, B, C vara händelser. Följande likheter gäller:
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- $\left(\bigcup_{i=1}^{n} A_i\right)^* = \bigcap_{i=1}^{n} A_i^*$
- $\left(\bigcap_{i=1}^{n} A_i\right)^* = \bigcup_{i=1}^{n} A_i^*$

Def (sannolikhet): Låt A vara en händelse. Det finns ett tal $P(A)$ som kallas sannolikheten för A. Man söker välja $P(A)$ så att den relativa frekvensen vid ett någorlunda stort antal försök kommer i närheten $P(A)$.

Def (frekvenstolkningen): Låt $a \in [0,1]$, om vi låter $P(A) = a$ kan vi ge detta uttalande den påtagliga men samtidigt vaga frekvenstolkningen: Vid ett stort antal försök blir den relativa frekvensen av händelsen A nog ungefär lika med a.

SATS 2.β (Kolmogrovs axiomsystem): Följande axiom för sannolikhetsmåttet $P(\cdot)$ skall vara uppfyllda:
- Axiom 1. För varje händelse A gäller att $0 \leq P(A) \leq 1$
- Axiom 2. För hela utfallsrummet Ω gäller att $P(\Omega) = 1$
- Axiom 3. Additionsformeln: Om $A_1, A_2, \ldots$ är en ändlig eller uppräckneligt oändlig följd av parvis oförenliga händelser gäller att
  $P(A_1 \cup A_2 \ldots) = P(A_1) + P(A_2) + \ldots$

Utfallsrummet Ω, händelserna $A, B, \ldots$ och sannolikheterna $P(\cdot)$ säges tillsammans utgöra ett sannolikhetsrum.

SATS 2.1 (Komplementsatsen): För komplementet A^* till A gäller att $P(A^*) = 1 - P(A)$

SATS 2.2 (Additionssatsen för två händelser): För två godtyckliga händelser A och B gäller att $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

SATS 2.3 (Booles olikhet): För två godtyckliga händelser A och B gäller att $P(A \cup B) \leq P(A) + P(B)$.

Def (likformigt sannolikhetsmått): Om $P(\omega_i) = 1/m$ för $i = 1, \ldots, m$ föreligger ett likformigt sannolikhetsmått.

SATS 2.4 (Den klassiska sannolikhetsdefinitionen): Vid likformigt sannolikhetsmått är sannolikheten för en händelse lika med kvoten mellan antalet för händelsens gynsamma fall och antalet möjliga fall.

SATS 2.8 (Multiplikationsprincipen): Om åtgärd 1 kan utföras på a_1 sätt och åtgärd 2 på a_2 sätt, så finns det $a_1 \cdot a_2$ sätt att utföra båda åtgärderna. Generalisering till tre åtgärder blir $a_1 a_2 a_3$.

SATS 2.5: Dragning med återläggning av k element ur n med hänsyn till ordning kan ske på n^k olika sätt

SATS 2.6: Dragning utan återläggning av k element ur n (med hänsyn till ordning) kan ske på $n(n-1)\ldots(n-k+1)$ olika sätt

Följdsats 2.6.1: Antalet permutationer av n element bland n är lika med $n(n-1)\ldots \cdot 2 \cdot 1 = n!$ eller kortare: n element kan ordnas på $n!$ olika sätt.

SATS 2.7: Dragning utan återläggning av k element ur n (utan hänsyn till ordning) kan ske på $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdots(n-k+1)}{k!}$ olika sätt

SATS 2.8 (Binomialteoremet): För varje positivt heltal n och för godtyckliga tal x och y gäller $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$

Def (betingad sannolikhet): Låt A och B vara två händelser. Uttrycket

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

kallas den betingande sannolikheten för B givet att A har inträffat. Om $P(A) = 0$ låter vi $P(B|A)$ vara obestämd.

Det gäller även att
- $P(B^*|A) = 1 - P(B|A)$
- $P(B \cup C|A) = P(B|A) + P(C|A) - P(B \cap C|A)$

Vi kan också skriva: $P(A \cap B) = P(A) \cdot P(B|A)$
$= P(B) \cdot P(A|B)$

SATS 2.9 (Lagen om total sannolikhet): Om händelserna $H_1, \ldots, H_n$ är parvis oförenliga och $\bigcap_{i=1}^{n} H_i = \Omega$, gäller för varje händelse A att $P(A) = \sum_{i=1}^{n} (P(H_i) P(A|H_i))$

SATS 2.10 (Bayes sats): Under samma villkor som i SATS 2.9 gäller

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{\sum_{j=1}^{n}(P(H_j)P(A|H_j))} = \frac{P(A \cap H_i)}{\sum_{j=1}^{n}(P(A \cap H_j))} = \frac{P(H_i) \cdot P(A|H_i)}{P(A)}$$

Def (oberoende): Om $P(A \cap B) = P(A) \cdot P(B)$ säges A och B vara oberoende händelser.

Def(oberoende): Om:

- $P(A \cap B) = P(A)P(B)$
- $P(A \cap C) = P(A)P(C)$
- $P(B \cap C) = P(B)P(C)$
- $P(A \cap B \cap C) = P(A)P(B)P(C)$

Säges A, B och C vara **oberoende** händelser
Om A och B är oberoende så är A^* och B oberoende.

SATS 2.11: Om händelserna $A_1, A_2, \dots, A_n$ är oberoende och $P(A_i) = p_i$, så är sannolikheten att minst en av dem inträffar $1-(1-p_1)(1-p_2)\dots(1-p_n)$

Följdsats 2.11.1: Om händelserna A_i är oberoende och var och en inträffar med sannolikheten p, så är sannolikheten att minst en av dem inträffar lika med $1-(1-p)^n$.

Def (stokastisk variabel): En **stokastisk variabel** (s.v) är en reellvärd funktion definierad på ett utfallsrum. $X: \Omega \to \mathbb{R}$

Def (diskret): En stokastisk variabel är **diskret** om den kan anta ett ändligt eller uppräkneligt oändligt antal olika värden.
Storheterna $p_X(x) = P(X = x)$, $x = a_1, a_2, \dots$
där $a_1, a_2, \dots$ är de (uppräkneligt många) tänkbara värdena som X kan anta, kallas **sannolikhetsfunktionen** för den s.v. X

SATS 3.α: Låt A vara en delmängd av de tänkbara värdena som X kan anta, d.v.s $A \subset V_X$. Vi har då: $P(X \in A) = \sum_{k \in A} p_X(k)$

Följdsats 3.α.1: Låt $V_X = \mathbb{N}$. Vi har då $P(X \in V_X) = \sum_{k=0}^{\infty} p_X(k) = 1$

Följdsats 3.α.2: Låt $V_X = \mathbb{N}$. Vi har då $P(a < X \leq b) = \sum_{k=a+1}^{b} p_X(k)$ där $a, b \in \mathbb{N}$

Def (tvåpunktsfördelad): Om den s.v. X antar endast två värdena a och b med sannolikheterna p respektive 1-p säges X vara **tvåpunktsfördelad**. Alltså $p_X(a) = p$, $p_X(b) = 1-p$.
I specialfallet då $a=1$ och $b=0$ säger man ofta att X är **Bernoulli-fördelad**.

Def (likformigt fördelad): Om den s.v. X antar värdena $1, 2, \dots, m$ med lika stor sannolikhet, $1/m$, säges X vara **likformigt fördelad**, eller mer specifikt har X en **likformig fördelning** över $\{1, 2, \dots, m\}$.

Def (för-första-gången-fördelad): Om den s.v. X har sannolikhetsfunktionen
$p_X(k) = (1-p)^{k-1} p$, $k = 1, 2, \dots$ där $p(0,1)$ säges X vara **för-första-gången-fördelad**. Betecknas $X \in ffg(p)$

Def (geometriskt fördelad): Om den s.v. X har sannolikhetsfunktionen $p_X(k) = (1-p)^k p$, $k = 0, 1, 2, \dots$
där $p \in (0,1)$, säges X vara **geometriskt fördelad**. Betecknas $X \in Ge(p)$

Def (binomialfördelad): Om den s.v. X har sannolikhetsfunktionen

$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, 2, \dots, n$ där n är ett positivt heltal och $p \in (0,1)$
säges X vara **binomialfördelad**. Betecknas $X \in Bin(n,p)$.
Enligt binomialtheoremet är
$\sum_{k=0}^{n} p_X(k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1^n = 1$

Def(hypergeometriskt fördelad): Om den s.v. X har sannolikhetsfunktionen

$$p_X(k)=\binom{v}{k}\binom{s}{n-k}\Big/\binom{v+s}{n}$$ där k antar alla heltalsvärden sådana att $0\le k\le v$, $0\le n-k\le s$ säges X vara hypergeometriskt fördelad. Betecknas $X\in Hyp(N,n,p)$ där $N=s+v$, och $p=\frac{v}{v+s}$

Def(Poisson-fördelning): Om den s.v. X har sannolikhetsfunktionen

$$p_X(k)=\frac{\mu^k}{k!}e^{-\mu},\quad k=0,1,2,\dots,\quad \mu>0$$ säges X vara Poisson-fördelad. Betecknas $X\in Po(\mu)$

Def(kontinuerlig): Om det finns en funktion $f_X(x)$ sådan att $P(X\in A)=\int_A f_X(x)\,dx$ för alla intervall $A\subset\mathbb{R}$, säges X vara en kontinuerlig s.v. Funktionen $f_X(x)$ kallas täthetsfunktionen för X. Kan även kallas frekvensfunktion.

SATS 3.1: I varje punkt x där $f_X(x)$ är kontinuerlig, gäller att
$$F_X'(x)=\frac{d}{dx}F_X(x)=f_X(x)$$

Def(likformigt fördelad): Om den s.v. X har täthetsfunktionen
$$f_X(x)=\begin{cases}1/(b-a) & \text{om } a<x<b\\ 0 & \text{annars}\end{cases}$$
säges X vara likformigt fördelad. Betecknas $X\in U(a,b)$.

Def(exponentialfördelad): Om den s.v. X har täthetsfunktionen
$$f_X(x)\begin{cases}\lambda e^{-\lambda x} & \text{om } x>0\\ 0 & \text{om } x\le 0\end{cases}$$
där $\lambda>0$ säges X vara exponentialfördelad. Betecknas $X\in Exp(\lambda)$

Def(normalfördelad): Om den s.v. X har täthetsfunktionen
$$f_X=\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/2\sigma^2},\quad x\in\mathbb{R}$$
där μ och σ är givna tal $(\sigma>0)$, säges X vara normalfördelad. Betecknas $X\in N(\mu,\sigma)$

Def(Weibull-fördelad): Om den s.v. X har täthetsfunktionen
$$f_X(x)=\begin{cases}\lambda c(\lambda x)^{c-1}e^{-(\lambda x)^c} & x>0\\ 0 & x\le 0\end{cases}$$
där λ och c är positiva tal, säges X vara Weibull-fördelad

Och
$$F_X(x)=\begin{cases}0 & \text{om } x<0\\ 1-e^{-(\lambda x)^c} & \text{om } x\ge 0\end{cases}$$

Def(gammafunktionen): Vi definierar gammafunktionen $\Gamma(c)$ som
$$\Gamma(c)=\int_0^\infty x^{c-1}e^{-x}\,dx.$$
Om c är ett heltal gäller $\Gamma(c)=(c-1)!$

Def(gammafördelad): Om den s.v. X har täthetsfunktionen
$$f_X(x)=\begin{cases}\frac{\lambda}{\Gamma(c)}\cdot x^{c-1}e^{-\lambda x} & \text{om } x>0\\ 0 & \text{om } x\le 0\end{cases}$$
där $c>0$ och $\lambda>0$, säges X vara gammafördelad

Def(fördelningsfunktion): $F_X(x)=P(X\le x)$, $x\in\mathbb{R}$, kallas fördelningsfunktionen för den s.v. X

SATS 3.β: För en kontinuerlig s.v. har vi: $F_X(x) = \int_{-\infty}^{x} f_X(t)dt = P(X \leq x)$

SATS 3.γ: För en diskret s.v har vi: $F_X(k) = \sum_{j \leq k} p_X(j)$ för $k = 0, 1, 2, \ldots$
och omvänt följer:
$$p_X(k) = \begin{cases} F_X(0) & \text{om } k=0 \\ F_X(k) - F_X(k-1) & \text{annars.} \end{cases}$$

SATS 3.2: För en fördelningsfunktion $F_X(x)$ gäller att

- $F_X(x) \to \begin{cases} 0 & \text{då } x \to -\infty \\ 1 & \text{då } x \to \infty \end{cases}$
- $F_X(x)$ är en växande funktion av x.
- $F_X(x)$ är kontinuerlig till höger för varje x

SATS 3.3: Om $a < b$ gäller att $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a) = \int_a^b f_X(t)dt$

Def(kvantil): Lösningen $x = x_\alpha$ till ekvationen $F_X(x) = 1 - \alpha$ kallas α-kvantilen för den s.v. X. α anges ofta i procent.

Def(kvartil, median): Kvantilerna $x_{0.25}$, $x_{0.5}$ och $x_{0.75}$ kallas övre kvartilen, medianen respektive undre kvartilen.

SATS 3.δ: Om X är en heltalsvariabel med sannolikhetsfunktionen $p_X(j)$ så har den s.v. $Y = g(X)$ sannolikhetsfunktionen

$$p_Y(k) = \sum_{j:g(j)=k} p_X(j)$$

SATS 3.ε: Om X är kontinuerlig och $Y = g(X)$ där g är strängt växande har vi
$F_Y(y) = F_X(g^{-1}(y))$
Om g är strängt avtagande har vi istället
$F_Y(y) = 1 - F_X(g^{-1}(y))$

Def(tvådimensionell stokastisk variabel): En tvådimensionell stokastisk variabel är en funktion (X, Y) definierad på ett utfallsrum Ω och som tar värden i $\mathbb{R}^2$.

Def(fördelningsfunktion): Funktionen $F_{X,Y}(x,y) = P(X \leq x, Y \leq y)$ kallas fördelningsfunktionen för den tvådimensionella s.v. (X, Y). Kan även kallas den simultana fördelningsfunktionen för X och Y.

Def(sannolikhetsfunktion): Storheterna $p_{X,Y}(j,k) = P(X=j, Y=k)$, $j = 0,1,2,\ldots$, $k = 0,1,2,\ldots$ kallas sannolikhetsfunktionen för den diskreta tvådimensionella s.v. (X, Y). Kan även mer specifikt kallas den simultana sannolikhetsfunktionen.

SATS 4.α: Sannolikheten $P((X,Y) \in A)$ kan beräknas genom $P((X,Y) \in A) = \sum_{(j,k) \in A} p_{X,Y}(x,y)$
Övrigt har vi att summan av alla sannolikheter är 1.
Alltså $\sum_{j=0} \sum_{k=0} p_{X,Y}(j,k) = 1$.

SATS 4.β: Fördelningsfunktionen för den diskreta tvådimensionella s.v. (X, Y) kan bestämmas som
$$F_{X,Y}(x,y) = \sum_{j \leq x} \sum_{k \leq y} p_{X,Y}(j,k).$$

Def(marginell sannolikhetsfunktion): För den diskreta tvådimensionella s.v. (X,Y) har vi att den marginella sannolikhetfunktionen för X i punkten j definieras som $p_X(j) = \sum_{k=0}^{\infty} p_{X,Y}(j,k)$
$p_Y(k)$ definieras på analogt vis.

Def (multinomialfördelning): Antag att ett slumpförsök har r olika utfall $A_1, A_2, \dots, A_r$ med sannolikheter $p_1, \dots, p_r$ där $p_1 + \dots + p_r = 1$. Låt X_j vara antalet gånger A_j inträffar vid n oberoende upprepningar av försöktet. Då gäller att $X_j \in \mathrm{Bin}(n, p_j)$. Då har vi att den simultana sannolikhetsfunktionen för den r-dimensionella s.v. $(X_1, \dots, X_r)$ ges av

$$p_{X_1,\dots,X_r}(k_1,\dots,k_r) \overset{*}{=} \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r} \quad \text{där } k_1,\dots,k_r \geq 0,\ k_1+\dots+k_r = n.$$

Vi säger att $(X_1, \dots, X_r)$ är multinomialfördelad. Lägg märke till $X_1 + \dots + X_r = n$

$$^{*}p_{X_1,\dots,X_r}(k_1,\dots,k_r) = \binom{n}{k_1} p_1^{k_1} \binom{n-k_1}{k_2} p_2^{k_2} \cdot \dots \cdot \binom{n-k_1-\dots-k_{r-1}}{k_r} p_r^{k_r} = \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r}$$

Def (täthetfunktion): Om det finns en funktion $f_{X,Y}(x,y)$ sådan att

$$P((X,Y) \in A) = \iint_A f_{X,Y}(x,y)\,dx\,dy.$$

för alla A säges den tvådimensionella s.v. (X,Y) vara kontinuerlig. Funktionen $f_{X,Y}(x,y)$ kallas täthetsfunktionen för (X,Y).
Vi har då

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v)\,du\,dv \quad \text{om } A \text{ är kvartsplanet snett ner till vänster från punkten } (x,y)$$

Som täthetsfunktion duger alla funktioner som uppfyller
$f(x,y) \geq 0$ för alla x,y och $\iint_{\mathbb{R}^2} f(x,y)\,dx\,dy = 1$.
$f_{X,Y}(x,y)$ kan kallas mer specifikt för den simultana täthetsfunktionen.
Vi har slutligen

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$$

Def (marginell täthetsfunktion, marginell fördelningsfunktion): Den marginella fördelningsfunktionen för X respektive Y ges av

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy \quad \text{och} \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx.$$

och den marginella fördelningsfunktionen för X respektive Y ges av

$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x,y) \quad \text{och} \quad F_Y(y) = \lim_{x \to \infty} F_{X,Y}(x,y)$$

Def (likformigt fördelad): En tvådimensionell s.v. (X,Y) är likformigt fördelad på området B, om för varje del A av detta område gäller att:

$$P((X,Y) \in A) = \frac{\text{Arean av } A}{\text{Arean av } B}$$

Def (oberoende): De s.v. X och Y kallas oberoende om $P(X \in C, Y \in D) = P(X \in C)P(Y \in D)$ för alla mängder C och D.

SATS 4.γ: Låt $Z = \max(X,Y)$ där X och Y är oberoende. Vi har då:

$$F_Z(z) = P(Z \leq z) = P(X \leq z \text{ och } Y \leq z) = P(X \leq z)P(Y \leq z) = F_X(z)F_Y(z)$$

SATS 4.δ: Låt $Z = \min(X,Y)$ där X och Y är oberoende. Vi har då:

$$F_Z(z) = P(Z \leq z) = 1 - P(Z > z) = 1 - P(X > z \text{ och } Y > z) = 1 - P(X > z)P(Y > z)$$

Alltså $F_Z(z) = 1 - (1 - F_X(z))(1 - F_Y(z))$

SATS 4.ε: Låt $Z_1 = \max(X_1, \dots, X_n)$ och $Z_2 = \min(X_1, \dots, X_n)$ där $X_1, \dots, X_n$ är n oberoende s.v. Vi har då $F_{Z_1} = \prod_{j=1}^{n} F_{X_j}(z)$ och $F_{Z_2} = 1 - \prod_{i=1}^{n} (1 - F_{X_i}(z))$

SATS 4.ζ: Låt X och Y vara diskreta s.v. som endast antar heltalsvärden och $Z=X+Y$.
Vi har då
$$p_Z(k)=P(X+Y=k)=\sum_{i+j=k} p_{X,Y}(i,j)=\sum_{i=0}^{k} p_{X,Y}(i,k-i)$$
och
$$F_Z(z)=\sum_{i+j\leq z} p_{X,Y}(i,j).$$

SATS 4.η: Låt X och Y vara kontinuerliga s.v. och $Z=X+Y$ har vi
$$F_Z(z)=P(X+Y\leq z)=\iint_{x+y\leq z} f_{X,Y}(x,y)\,dx\,dy$$
Om X och Y är oberoende har vi då
$$F_Z(z)=\iint_{x+y\leq z} f_X(x)f_Y(y)\,dx\,dy=\int_{-\infty}^{\infty} f_X(x)\left(\int_{-\infty}^{z-x} f_Y(y)\,dy\right)dx=\int_{-\infty}^{\infty} f_X(x)F_Y(z-x)\,dx$$

SATS 4.θ (faltningsformel): Låt X och Y vara oberoende och $Z=X+Y$. Då är faltningsformeln:

- $p_Z(k)=\sum_{i=0}^{k} p_X(i)p_Y(k-i)$ då X och Y är diskreta s.v.
- $f_Z(z)=\int_{-\infty}^{\infty} f_X(x)f_Y(z-x)\,dx$ då X och Y är kontinuerliga.

Def (väntevärde): Väntevärdet för den s.v. X definieras av
$$E(X)=\begin{cases}\sum_k k\,p_X(k) & \text{om } X \text{ är diskret}\\ \int_{-\infty}^{\infty} x f_X(x)\,dx & \text{om } X \text{ är kontinuerlig}\end{cases}$$
Ibland kallas E även för förväntat värde.

SATS 5.1: Om $Y=g(X)$ gäller det att
$$E(Y)=\begin{cases}\sum_k g(k)\,p_X(k) & \text{då } Y \text{ och } X \text{ är diskreta}\\ \int_{-\infty}^{\infty} g(x)f_X(x)\,dx & \text{då } Y \text{ och } X \text{ är kontinuerliga.}\end{cases}$$

Följdsats 5.1.1: Om $Y=aX+b$ har vi att $E(Y)=aE(X)+b$.

SATS 5.2: Om $Z=g(X,Y)$ gäller det att
$$E(Z)=\begin{cases}\sum_{j,k} g(j,k)\,p_{X,Y}(j,k) & \text{för diskreta s.v.}\\ \iint_{\mathbb{R}^2} g(x,y)\,f_{X,Y}(x,y)\,dx\,dy & \text{för kontinuerliga s.v.}\end{cases}$$

SATS 5.3: Om X och Y är s.v. med väntevärden $E(X)$ och $E(Y)$ och om $a,b,c\in\mathbb{R}$ så gäller det att $E(aX+bY+c)=aE(X)+bE(Y)+c$

SATS 5.4: Om X och Y är oberoende gäller det att $E(XY)=E(X)E(Y)$

SATS 5.5: Om $X_1,X_2,\dots,X_n$ är oberoende s.v. gäller $E(X_1X_2\dots X_n)=E(X_1)E(X_2)\dots E(X_n)$

Def (varians): Variansen $V(X)$ för en s.v. X med $E(X)=\mu$ är $V(X)=E\left((X-\mu)^2\right)$
Alltså
$$V(X)=\begin{cases}\sum_k (k-\mu)^2 p_X(k) & \text{då } X \text{ är en diskret s.v.}\\ \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x)\,dx & \text{då } X \text{ är en kontinuerlig s.v.}\end{cases}$$

Def (standardavvikelse): Standardavvikelsen $D(X)$ för en s.v. X är kvadratroten ur variansen $D(X)=\sqrt{V(X)}$

Def (variationskoefficient): Kvoten $R(X)=D(X)/E(X)$ kallas variationskoefficienten.

SATS 5.6: Följande samband gäller: $V(X) = E(X^2) - (E(X))^2$

SATS 5.7: Vid linjärtransformation gäller

$$E(aX+b) = aE(X)+b$$
$$V(aX+b) = a^2 V(X)$$
$$D(aX+b) = |a| D(X)$$

Def(standardiserad): Om X är en s.v. med väntevärdet $E(X)=\mu$ och standardavvikelsen $D(X)=\sigma$ kallas $Y=(X-\mu)/\sigma$ en standardiserad s.v. vilket har egenskaperna

$$\begin{cases} E(Y)=0 \\ D(Y)=1 \end{cases}$$

Def(systematiskt fel, slumpmässigt fel): Med systematiskt fel menas differensen mellan mätvärdets väntevärde och det korrekta värdet. Med slumpmässigt fel menas differensen mellan mätvärdet och dess väntevärde.

Def(noggranhet, precision): Med god noggranhet avses ett litet systematiskt fel. Med god precision avses ett litet slumpmässigt fel.

Def(kovarians): Kovariansen $C(X,Y)$ mellan X och Y är

$$C(X,Y) = E((X-\mu_X)(Y-\mu_Y))$$

Alltså från sats 5.2:

$$C(X,Y) = \begin{cases} \sum_j \sum_k (j-\mu_X)(k-\mu_Y) p_{X,Y}(j,k) \\ \iint_{\mathbb{R}^2} (x-\mu_X)(y-\mu_Y) f_{X,Y}(x,y)\,dx\,dy \end{cases}$$

SATS 5.8: Följande samband gäller:

$$C(X,Y) = E(XY) - E(X)E(Y) = E(XY) - \mu_X \mu_Y \quad \text{där } E(XY) = \begin{cases} \sum_{j,k} j \cdot k \cdot p_{X,Y}(k) \\ \iint_{\mathbb{R}^2} xy \cdot f_{X,Y}(x,y)\,dx\,dy \end{cases}$$

SATS 5.α: Kovariansen bör vara positiv om det finns ett beroende mellan X och Y sådant att det finns en tendens hos variablerna att samtidigt avvika åt samma håll från sina respektive väntevärden. Om variablerna däremot tenderar att avvika åt olika håll från väntevärdena bör kovariansen vara negativ.
Detta påstående kommer från undersökning av produkten $(X-\mu_X)(Y-\mu_Y)$.

Def(korrelationskoefficienten): Korrelationskoefficienten för X och Y definieras av

$$\rho(X,Y) = \frac{C(X,Y)}{D(X)D(Y)}$$

Def(okorrelerad): Om $C(X,Y)=0$ säges X och Y vara okorrelerade.

SATS 5.9: Om X och Y är oberoende så är de också okorrelerade.
OBS: Att X och Y är okorrelerade innebär inte att de är oberoende!

SATS 5.β: För alla s.v. X och Y har vi: $-1 \leq \rho(X,Y) \leq 1$

Om $\rho(X,Y)$ antar ett:

- Negativt värde nära -1 är X och Y starkt negativt korrelerade
- Negativt värde nära 0 är X och Y svagt negativt korrelerade
- Positivt värde nära 0 är X och Y svagt positivt korrelerade
- Positivt värde nära 1 är X och Y starkt positivt korrelerade

SATS 5.10: För alla s.v. X och Y gäller att

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

$V(X+Y) = V(X) + V(Y) + 2C(X,Y)$

Följdsats 5.10.1: För oberoende s.v. X och Y gäller att

$V(X+Y) = V(X) + V(Y)$

$D(X+Y) = \sqrt{D^2(X) + D^2(Y)}$

$V(X-Y) = V(X+Y)$

SATS 5.11: För alla s.v. $X_1, \ldots, X_n$ gäller att

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

och

$$V\left(\left(\sum_{i=1}^{n} a_i X_i\right) + b\right) = \sum_{i=1}^{n} a_i^2 V(X_i) + 2 \sum_{1 \le j < k \le n} a_j a_k C(X_j, X_k)$$

Följdsats 5.11.1: För oberoende s.v. $X_1, \ldots, X_n$ gäller att

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

och

$$V\left(\left(\sum_{i=1}^{n} a_i X_i\right) + b\right) = \sum_{i=1}^{n} a_i^2 V(X_i)$$

Följdsats 5.11.2: Om $X_1, \ldots, X_n$ är en s.v. med samma väntevärde μ så gäller att

$$E\left(\sum_{i=1}^{n} X_i\right) = n\mu$$

Och om $X_1, \ldots, X_n$ är oberoende och har samma standardavvikelse σ gäller även

$$V\left(\sum_{i=1}^{n} X_i\right) = n\sigma^2 \quad \text{och} \quad D\left(\sum_{i=1}^{n} X_i\right) = \sigma\sqrt{n}$$

Följdsats 5.11.3: Om $X_1, \ldots, X_n$ är oberoende s.v. var och en med väntevärde μ och standardavvikelsen σ och

$\overline{X} = \sum_{i=1}^{n} X_i / n$ är deras arithmetiska medelvärde, så gäller att

$$E(\overline{X}) = \mu, \quad V(\overline{X}) = \sigma^2/n \quad \text{och} \quad D(\overline{X}) = \sigma/\sqrt{n}$$

SATS 5.12 (Stora talens lag): Låt $X_1, X_2, \ldots$ vara oberoende och likafördelade s.v., var och en med väntevärdet μ, och sätt

$\overline{X}_n = \sum_{i=1}^{n} X_i / n$

Då gäller, för alla $\varepsilon > 0$, att

$P(\mu - \varepsilon < \overline{X}_n < \mu + \varepsilon) \to 1$ då $n \to \infty$

SATS 5.13 (Markovs olikhet): För $a > 0$ och $Y \ge 0$ gäller $P(Y \ge a) \le \frac{E(Y)}{a}$

SATS 5.14 (Tjebysjovs olikhet): Låt X vara en s.v. med väntevärde μ och standardavvikelse $\sigma > 0$. Då gäller, för varje $k > 0$, att $P(|X - \mu| \ge k\sigma) \le 1/k^2$

Def(standardiserad normalfördelad): Låt $X \in N(0,1)$. Vi säger då att X är en **standardiserad normalfördelad** s.v. Täthetsfunktionen för denna s.v. brukar betecknas $\varphi(x)$ och fördelningsfunktionen $\Phi(x)$. Man har följdaktligen:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2} \quad ; \quad \Phi(x) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$

SATS 6.α: Låt $X \in N(0,1)$. Vi har då $\varphi(-x) = \varphi(x)$ och $\Phi(-x) = 1 - (\Phi(x))$.

Def(kvantil): **α-kvantilen** för en standardiserad normalfördelning betecknas λ_α. D.v.s. $P(X > \lambda_\alpha) = \alpha$. Vi har även $\lambda_{1-\alpha} = -\lambda_\alpha$ p.g.a. symmetri kring 0.

SATS 6.β: Låt $X \in N(0,1)$. Då har vi $P(-\lambda_\alpha < X < \lambda_\alpha) = 1 - 2\alpha$

SATS 6.γ: Låt $X \in N(0,1)$. Då har vi $E(X) = 0$ och $D(X) = 1$

SATS 6.1: $X \in N(\mu, \sigma)$ om och endast om $Y = (X-\mu)/\sigma \in N(0,1)$. Dessutom gäller att

$$f_X(x) = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right) \quad \text{och} \quad F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

SATS 6.2: Om $X \in N(\mu, \sigma)$, så gäller att
$E(X) = \mu$, $V(X) = \sigma^2$, och $D(X) = \sigma$

SATS 6.δ: Om $X \in N(\mu, \sigma)$ så gäller att $P(\mu + k_1\sigma < X < \mu + k_2\sigma) = \Phi(k_2) - \Phi(k_1)$

SATS 6.3: Om $X \in N(\mu, \sigma)$, så gäller att $Y = aX + b \in N(a\mu + b, |a|\sigma)$

SATS 6.4: Om $X \in N(\mu_X, \sigma_X)$ och $Y \in N(\mu_Y, \sigma_Y)$, där X och Y är oberoende, så gäller att

$$X + Y \in N\left(\mu_X + \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2}\right)$$

$$X - Y \in N\left(\mu_X - \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2}\right)$$

SATS 6.5: Om $X_1, X_2, \ldots, X_n$ är oberoende och respektive $N(\mu_1, \sigma_1), \ldots, N(\mu_n, \sigma_n)$ och talen $a_1, \ldots, a_n$ och b är givna, så gäller
$\sum_1^n a_i X_i + b \in N\left(\sum_1^n a_i\mu_i + b, \sqrt{\sum_1^n a_i^2 \sigma_i^2}\right)$

Följdsats 6.5.1: Om $X_1, X_2, \ldots, X_n$ är oberoende $N(\mu, \sigma)$ och $\bar{X} = \sum_1^n X_i/n$ är deras aritmetiska medelvärde, så gäller att $\bar{X} \in N(\mu, \sigma/\sqrt{n})$

Följdsats 6.5.2: Om $X_1, X_2, \ldots, X_{n_1}$ är $N(\mu_1, \sigma_1)$ och $Y_1, Y_2, \ldots, Y_{n_2}$ är $N(\mu_2, \sigma_2)$ och alla variabler är oberoende, så gäller att

$$\bar{X} - \bar{Y} \in N\left(\mu_1 - \mu_2, \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}\right)$$

Def(χ^2-fördelad): Om den s.v. X har täthetsfunktionen

$$f_X(x) = \begin{cases} \dfrac{x^{f/2-1} e^{-x/2}}{\Gamma(f/2)2^{f/2}} & \text{om } x > 0 \\ 0 & \text{om } x \leq 0 \end{cases}$$

säges X vara **χ^2-fördelad** med f frihetsgrader. Betecknas $X \in \chi^2(f)$

SATS 6.6: Om $X_1, X_2, \ldots, X_f$ är oberoende och $N(0,1)$, så gäller att $\sum_{i=1}^{f} X_i^2 \in \chi^2(f)$.

SATS 6.8 (Centrala gränsvärdessatsen): Om $X_1, X_2, \dots$ är en oändlig följd av oberoende likafördelade s.v med μ och standardavvikelsen $\sigma > 0$, så gäller för $Y_n = \sum_{i=1}^{n} X_i$ att

$$P\left(a < \frac{Y_n - n\mu}{\sigma\sqrt{n}} \leq b\right) \to \Phi(b) - \Phi(a) \text{ då } n \to \infty.$$

Def (asymptotiskt normalfördelad): Om Z_n, $n = 1, 2, \dots$, är en oändlig följd av s.v. och man kan finna tal A_n och B_n, $n = 1, 2, \dots$ sådana att

$$P\left(a < \frac{Z_n - A_n}{B_n} \leq b\right) \to \Phi(b) - \Phi(a) \text{ då } n \to \infty$$

sägs Z_n vara asymptotiskt normalfördelad med parametrarna A_n och B_n.
Betecknas $Z_n \in AsN(A_n, B_n)$

Notera: SATS 6.8 säger att $Y_n \in AsN(n\mu, \sigma\sqrt{n})$

Följdsats 6.8.1: Om $X_1, X_2, \dots$ är en följd av oberoende likafördelade s.v. med väntevärde μ och standardavvikelsen $\sigma > 0$ så gäller att $(X_1 + \dots + X_n)/n \in AsN(\mu, \sigma/\sqrt{n})$

SATS 7.1: Antag att sannolikheten är p att en händelse A inträffar i ett enskilt försök. Om n oberoende försök utförs och X är antalet gånger som A inträffar, gäller det att $X \in Bin(n, p)$.

Def (indikatorvariabel): Antag att sannolikheten är p att en händelse inträffar i ett enskilt försök. Om n oberoende försök utförs och I_j är en s.v. som är 1 om A inträffade på j:te försöket och annars är noll för $j = 1, 2, 3 \dots, n$ kallar vi $I_1, I_2, \dots, I_n$ för indikatorvariabler och $X = I_1 + I_2 + \dots + I_n \in Bin(n, p)$.

SATS 7.2: Om X är $Bin(n, p)$ gäller det att
$E(X) = np \quad V(X) = np(1-p) \quad D(X) = \sqrt{np(1-p)}$

SATS 7.3: Om $X \in Bin(n_1, p)$ och $Y \in Bin(n_2, p)$ där X och Y är oberoende, gäller det att $X + Y \in Bin(n_1 + n_2, p)$

SATS 7.α: Om $X \in Bin(n, p)$ kan $P(a < X \leq b)$ approximeras som

$$P(a < X \leq b) \approx \Phi\left(\frac{b - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - np}{\sqrt{np(1-p)}}\right) \quad \text{för } np(1-p) \geq 10 \text{ och } a, b \in \mathbb{N}$$

eller, för mer exakta värden:

$$P(a < X \leq b) \approx \Phi\left(\frac{b + 1/2 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a + 1/2 - np}{\sqrt{np(1-p)}}\right) \quad \text{för } np(1-p) \geq 3 \text{ och } a, b \in \mathbb{N}$$

Alltså
- $Bin(n,p) \sim N(np, \sqrt{np(1-p)})$: $np(1-p) \geq 10$
- $Bin(n,p) \sim N(np - 1/2, \sqrt{np(1-p)})$: $np(1-p) \geq 3$

SATS 7.β: Om $X \in Bin(n, p)$ där $p \leq 0.1$ kan $p_X(k)$ approximeras som
$p_X(k) \approx \frac{(np)^k}{k!} e^{-np}$, $k = 1, 2, \dots$
Alltså
$Bin(n,p) \sim Po(np)$: $p \leq 0.1$

Def (relativ frekvens): I vart och ett av n oberoende försök noteras om en händelse A inträffar. Vi låter X vara antalet gånger A inträffade, alltså $X \in Bin(n, p)$ där p är sannolikheten att A inträffar på ett försök. $Y = X/n$ kallas relativa frekvensen för A. Vi har $0 \leq Y \leq 1$ och
$E(Y) = p$, $V(Y) = p(1-p)/n$, $D(Y) = \sqrt{p(1-p)/n}$

SATS 7.4 (Bernoullis sats): Sätt $Y=X/n$ där $X \in Bin(n,p)$. För varje $\varepsilon>0$ gäller då att $P(|Y-p|>\varepsilon) \to 0$ då $n \to \infty$.

SATS 7.γ: Sätt $Y=X/n$ där $X \in Bin(n,p)$. För stora n har vi att Y kan approximeras som $N(p, \sqrt{p(1-p)/n})$

SATS 7.5: Låt relativa frekvensen av A-element i en population av N element vara p. Om n element uttas slumpmässigt utan återläggning, är antalet A-element X bland dessa hypergeometriskt fördelat:

$$p_X(k) = \frac{\binom{Np}{k}\binom{N(1-p)}{n-k}}{\binom{N}{n}}$$

k antar alla heltal sådanna att $0 \leq k \leq Np$, $0 \leq n-k \leq N(1-p)$

SATS 7.6: För en hypergeometriskt fördelad s.v. X gäller att

$$E(X) = np, \quad V(X) = \frac{N-n}{N-1} \cdot np(1-p)$$

Def (korrektionsfaktor): Låt $d_n^2 = \frac{N-n}{N-1}$ och $d_n = \sqrt{\frac{N-n}{N-1}}$. Vi kallar d_n^2 för korrektionsfaktorn för ändlig population.

SATS 7.δ: Om $X \in Hyp(N,n,p)$ kan X approximeras som

- $X \in Bin(n,p)$ om $\frac{n}{N} < 0.1$
- $X \in N(np, \sqrt{\frac{N-n}{N-1} \cdot np(1-p)})$ om $\frac{N-n}{N-1} \cdot np(1-p) \geq 10$
- $X \in Po(np)$ om $p + (n/N) < 0.1$

SATS 7.ε: Låt $Y=X/n$. Vi har då:

$$E(Y) = p, \quad V(Y) = d_n^2\, p(1-p)/n, \quad D(Y) = d_n \sqrt{p(1-p)/n}$$

SATS 7.ζ: Poisson-fördelningen uppträder då man studerar företeelser som inträffar slumpmässigt i tiden eller rummet, oberoende av varandra, d.v.s händelser som dels kan inträffa vilken tidpunkt som helst och är oberoende av varandra. Vi antar även att händelserna inträffar med konstant intensitet, så att λ händelser inträffar i genomsnitt per tids enhet, d.v.s. λt händelser inträffar i genomsnitt under en tidsperiod av längden t. I så fall gäller att $X \in Po(\lambda t)$

SATS 7.7: Om X är $Po(\mu)$ gäller det att $E(X) = \mu$, $V(X) = \mu$, $D(X) = \sqrt{\mu}$

SATS 7.8: Om $X \in Po(\mu_1)$ och $Y \in Po(\mu_2)$, där X och Y är oberoende, gäller det att $X+Y \in Po(\mu_1+\mu_2)$

SATS 7.η: Om $X \in Po(\mu)$ och $\mu > 15$ kan X approximeras som $X \in N(\mu, \sqrt{\mu})$

Def (Ogrupperad data): En mängd data som ej blivit sorterad kallas ogrupperad data.

Def (Gruppering): Gruppering av data är att sammanföra data av samma värde. Resultatet kallas grupperad data och kan representeras i t.ex. en frekvenstabell eller ett stolpdiagram.

Def(klassindelning): Klassindelning av data är att sammanföra data av ungefär samma storlek. En viss datapunkt a läggs i den i:te klassen om $a \in [g_i, g_{i+1})$. g_i och g_{i+1} kallas klassgränser och klassbredden $h = g_{i+1} - g_i$ bör vara konstant för alla definierade i.

Def(klassmitt): Värdet för den i:te klassen i en mängd klassindelad data kan ofta representeras av klassens klassmitt $y_i = (g_i + g_{i+1})/2$

Def(aritmetiskt medelvärde): Låt $x_1, x_2, \ldots, x_n$ vara en mängd ogrupperad data. Det aritmetiska medelvärdet $\bar{x}$ beräknas som:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Def(lägesmått, spridninsmått): Ett lägesmått är ett värde som ska representera hur stor kapacitansen är i "medeltal", ofta används ett medelvärde eller medianen. Ett spridningsmått är ett värde som på ett lämpligt sätt anger hur mycket de olika kapacitanserna skiljer sig åt från varandra, ofta används (stickprovs)variansen eller (stickprovets) standardavvikelse, se nedan definitioner.

Def(varians): (Stickprovs)variansen s^2 för en mängd data $x_1, x_2, \ldots, x_n$ definieras som

$$s^2 = \frac{1}{n-1}\sum_{j=1}^{n} (x_j - \bar{x})^2$$

Def(standardavvikelse): (Stickprovets) standardavvikelse s för en mängd data $x_1, \ldots, x_n$ definieras som

$$s = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n} (x_j - \bar{x})^2}$$

Alltså kvadratroten av variansen.

Def(kvadratsumma): Kvadratsumman kring det aritmetiska medelvärdet definieras som

$$Q = \sum_{j=1}^{n} (x_j - \bar{x})^2 = (n-1)s^2$$

eller (likvärdigt)

$$Q = \sum_{j=1}^{n} x_j^2 - \frac{1}{n}\left(\sum_{j=1}^{n} x_j\right)^2$$

Def(variationskoefficient): Variationskoefficienten är värdet $(100 \cdot \frac{s}{\bar{x}})\%$

Def(kvartilavstånd, variationsbredd): Låt övre och undre kvartilen av en datagrupp vara $\tilde{x}_{övre}$ respektive $\tilde{x}_{undre}$. Kvartilavståndet definieras av $\tilde{x}_{övre} - \tilde{x}_{undre}$ och kvartilintervallet definieras som $(\tilde{x}_{undre}, \tilde{x}_{övre})$. Variationsbredden definieras av $R = x_{max} - x_{min}$.

SATS 10.α: Om man har en mängd grupperad data av klasser där den i:te klassen har värde y_i och absolutfrekvensen f_i. Vi har då:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} f_i y_i, \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{k} f_i (y_i - \bar{x})^2 \quad \text{och} \quad Q = \sum_{i=1}^{k} f_i (y_i - \bar{x})^2 = \sum_{i=1}^{k} f_i y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{k} f_i y_i\right)^2$$

Ovanstående ekvationer gäller även approximativt för klassindelad data med klassmitt y_i.

Def(kovarians, korrelationskoefficient): Med kovariansen mellan x- och y-värdena i en datamängd $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ menas $c_{xy} = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$, och med korrelationskoefficienten menas $r = \frac{c_{xy}}{s_x s_y}$, där s_x och s_y är stickprovsstandardavvikelserna för x- respektive y-data.

Def(punktskattning, stickprovsvariabeln): En punktskattning $\theta^*_{obs} = \theta^*(x_1, x_2, \ldots, x_n)$ av en parameter θ är en funktion av mätdata $x_1, x_2, \ldots, x_n$. Dessa mätdata ses som utfall av stokastiska variabler $X_1, X_2, \ldots, X_n$ vilkas fördelning beror av parametern θ. Punktskattningen θ^*_{obs} är ett utfall av stickprovsvariabeln $\theta^* = \theta^*(X_1, X_2, \ldots, X_n)$
OBS: θ^*_{obs} är ett tal, θ^* är en stokastisk variabel. θ^*_{obs} är ett utfall av θ^*

Def(väntevärdesriktig): En punktskattning θ^*_{obs} säges vara väntevärdesriktig om tillhörande stickprovsvariabel θ^* har väntevärdet θ.
d.v.s. om
$$E(\theta^*) = \theta$$
för varje $\theta \in \Omega_\Theta$. (Ω_Θ är parameterrummet av θ)

Def(konsistent): Låt θ^*_n beteckna θ^*_{obs} för n mätdata. Om för varje fixerat $\theta \in \Omega_\Theta$ och för varje givet $\varepsilon > 0$ $P(|\theta^*_n - \theta| > \varepsilon) \to 0$ då stickprovsstorleken $n \to \infty$, säges θ^*_{obs} vara konsistent.

Def(medelkvadratfel): Medelkvadratfelet MSE för en punktskattning θ^*_{obs} är
$MSE = E((\theta^* - \theta)^2)$ där θ^* är motsvarande stickprovsvariabel
$MSE = E((\theta^* - \theta)^2) = E((\theta^* - E(\theta^*) + E(\theta^*) - \theta)^2) = V(\theta^*) + (E(\theta^*) - \theta)^2$

Def(systematiskt fel/bias): $E(\theta^*) - \theta$ brukar kallas systematiskt fel eller bias. Att det systematiska felet är 0 är ekvivalent med att skattningen är väntevärdesriktig.

Def(effektivare): Om två skattningar θ^*_{obs} och $\hat{\theta}_{obs}$ är väntevärdesriktiga och motsvarande stickprovsvariabler uppfyller $V(\theta^*) \leq V(\hat{\theta})$ för alla $\theta \in \Omega_\Theta$ (med sträng olikhet för något $\theta \in \Omega_\Theta$), så är θ^*_{obs} effektivare än $\hat{\theta}_{obs}$. $V(\theta^*)/V(\hat{\theta})$ kallas effektiviteten hos den sämre skattningen relativt den bättre.

SATS 11.1: Stickprovsmedelvärdet $\bar{x}$ är en väntevärdesriktig och konsistent skattning av μ.

SATS 11.2: Stickprovsvariansen s^2 är en väntevärdesriktig skattning av σ^2.

Def(likelihood-funktionen): Funktionen
$$L(\theta) = \begin{cases} P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n; \theta) & \text{(diskreta fallet)} \\ f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n; \theta) \end{cases}$$
kallas likelihood-funktionen (förkortat: L-funktionen)

Def(ML-skattning): Det värde θ^*_{obs}, för vilket $L(\theta)$ antar sitt största värde inom Ω_Θ, kallas ML-skattningen av θ
OBS: ML-skattningen är inte alltid väntevärdesriktig men kan ofta lätt korrigeras till att bli det.

Def(MK-skattningen): Låt $Q(\theta) = \sum_{i=1}^{n} (x_i - \mu_i(\theta))^2$ där $\mu_i(\theta)$ är väntevärdet för X_i.
Det värde θ^*_{obs}, för vilket $Q(\theta)$ antar sitt minsta värde inom Ω_Θ, kallas MK-skattningen av θ
Om $\forall i$ $\mu_i(\theta) = \mu(\theta)$ erhåller man $dQ(\theta)/d\theta = -2\mu'(\theta)\sum_1^n [x_i - \mu(\theta)]$ som visar att $Q(\theta)$ minimeras om $\mu(\theta) = \bar{x}$ som ger $\theta^*_{obs} = \mu^{-1}(\bar{x})$.
För k okända parametrar minimeras istället $Q(\theta_1, \theta_2, \ldots, \theta_k) = \sum_{i=1}^{n} (x_i - \mu_i(\theta_1, \theta_2, \ldots, \theta_k))^2$ för att bestämma $(\theta_1)^*_{obs}, (\theta_2)^*_{obs}, \ldots, (\theta_k)^*_{obs}$.

Def(medelfel): En skattning av $D(\theta^*)$ kallas medelfelet för θ^* och betecknas $d(\theta^*)$

Def (konfidensintervall, konfidensgrad, konfidensgräns): Ett intervall I_θ som med sannolikheten $1-\alpha$ innehåller θ kallas ett konfidensintervall för θ med konfidensgraden $1-\alpha$. Låt $\bar{x}=(x_1,\dots,x_n)$ vara det slumpmässiga stickprov som används för att göra en punktskattning på I_θ. Intervallets vänstra och högra ändpunkt kallas konfidensgränserna och betecknas allmänt $a_1(\bar{x})$ och $a_2(\bar{x})$. $a_1(\bar{x})$ och $a_2(\bar{x})$ är funktioner av värdena i stickprovet och alltså observationer av stickprovsvariabler $a_1(\bar{X})$ och $a_2(\bar{X})$
Definitionen ger oss:

$$P(\theta \in I_\theta) = P(a_1(\bar{X})<\theta<a_2(\bar{X}))=1-\alpha$$

Ett konfidensintervall $(a_1(\bar{x}), a_2(\bar{x}))$ kan sägas vara en observation av ett intervall med stokastiska gränser.

Def (tvåsidig, ensidig): Ett konfidensintervall $(a_1(\bar{x}), a_2(\bar{x}))$ kallas tvåsidigt om båda gränserna $a_1(\bar{x})$ och $a_2(\bar{x})$ är ändliga. Konfidensintervallen $(a_1(\bar{x}), \infty)$ och $(-\infty, a_2(\bar{x}))$ kallas ensidiga.

SATS 12.1: Låt $x_1,\dots,x_n$ vara ett slumpmässigt stickprov från $N(\mu,\sigma)$ där μ är okänt. Då är

$$I_\mu=(\bar{x}-\lambda_{\alpha/2}D,\ \bar{x}+\lambda_{\alpha/2}D) \quad \text{om } \sigma \text{ är känt och } D=\sigma/\sqrt{n}$$

samt

$$I_\mu=(\bar{x}-t_{\alpha/2}(f)d,\ \bar{x}+t_{\alpha/2}(f)d) \quad \text{om } \sigma \text{ är okänt, } d=s/\sqrt{n} \text{ och } f=n-1$$

ett tvåsidigt konfidensintervall för μ med konfidensgraden $1-\alpha$

SATS 12.2: Låt $x_1,\dots,x_n$ vara ett slumpmässigt stickprov från $N(\mu,\sigma)$. Då är $I_\sigma=(k_1 s, k_2 s)$ där

$$k_1=\sqrt{f/\chi^2_{\alpha/2}(f)} \quad \text{och} \quad k_2=\sqrt{f/\chi^2_{1-\alpha/2}(f)} \quad \text{för } f=n-1$$

ett konfidensintervall för σ med konfidensgraden $1-\alpha$.

SATS 12.3: Låt $x_1,\dots,x_{n_1}$ och $y_1,\dots,y_{n_2}$ vara slumpmässiga, av varandra oberoende stickprov från $N(\mu_1,\sigma_1)$ respektive $N(\mu_2,\sigma_2)$.
Om σ_1 och σ_2 är kända så är

$$I_{\mu_1-\mu_2}=(\bar{x}-\bar{y}-\lambda_{\alpha/2}D,\ \bar{x}-\bar{y}+\lambda_{\alpha/2}D)$$

ett tvåsidigt konfidensintervall för $\mu_1-\mu_2$ med konfidensgraden $1-\alpha$; här är $D=\sqrt{\sigma_1^2/n_1+\sigma_2^2/n_2}$.
Om $\sigma_1=\sigma_2=\sigma$ där σ är okänt så är

$$I_{\mu_1-\mu_2}=(\bar{x}-\bar{y}-t_{\alpha/2}(f)d,\ \bar{x}-\bar{y}+t_{\alpha/2}(f)d)$$

ett tvåsidigt konfidensintervall för $\mu_1-\mu_2$ med konfidensgraden $1-\alpha$; här är $d=s\sqrt{1/n_1+1/n_2}$ där $s=\sqrt{\frac{Q_1+Q_2}{(n_1-1)+(n_2-1)}}$ och $f=(n_1-1)+(n_2-1)$.

SATS 12.α: Låt A och B vara två olika sätt att mäta n objek där x_j och y_j är måttet på det j:te objektet med mättsätt A respektive B. Om normalfördelning föreligger antas x_j komma från en normalfördelning $N(\mu_j,\sigma_1)$ och y_j från en normalfördelning $N(\mu_j+\Delta,\sigma_2)$ där Δ anger den systematiska skillnaden mellan B:s och A:s mätningsätt (Om $\Delta>0$ är B:s värden i genomsnitt större än A:s). Om vi låter $z_1=y_1-x_1,\dots,z_n=y_n-x_n$ ser vi att $z_1,\dots,z_n$ är stickprov tagna från $N(\Delta,\sigma)$ där $\sigma=\sqrt{\sigma_1^2+\sigma_2^2}$. Konfidensintervall för Δ och σ kan då hittas med hjälp av SATS 12.1 respektive SATS 12.2

SATS 12.4: Antag att man på grundval av ett eller flera slumpmässiga stickprov har beräknat en punktskattning θ^* av en okänd parameter θ. Antag vidare att denna skattning är ungefär normalfördelad med väntevärdet θ och standardavvikelsen D. Då är:

$$I_\theta = (\theta^*_{obs} - \lambda_{\alpha/2} D,\ \theta^*_{obs} + \lambda_{\alpha/2} D) \quad \text{om } D \text{ ej beror av } \theta$$
$$I_\theta = (\theta^*_{obs} - \lambda_{\alpha/2} d,\ \theta^*_{obs} + \lambda_{\alpha/2} d) \quad \text{om } D \text{ beror av } \theta$$

(och d väljs lämpligt) ett konfidensintervall för θ med den approximativa konfidensgraden $1-\alpha$

Följdsats 12.4.1: Låt $x_1, \ldots, x_n$ vara ett stort slumpmässigt stickprov från en fördelning där väntevärdet μ och standardavvikelsen σ. Då är

$$I_\mu = (\bar{x} - \lambda_{\alpha/2} D,\ \bar{x} + \lambda_{\alpha/2} D) \quad \text{om } \sigma \text{ är känt och } D = \sigma/\sqrt{n}$$
$$I_\mu = (\bar{x} - \lambda_{\alpha/2} d,\ \bar{x} + \lambda_{\alpha/2} d) \quad \text{om } \sigma \text{ är okänt och } d = s/\sqrt{n}$$

konfidensintervall för μ med den approximativa konfidensgraden $1-\alpha$.

Följdsats 12.4.2: Låt $x_1, \ldots, x_{n_1}$ och $y_1, \ldots, y_{n_2}$ vara slumpmässiga och av varandra oberoende stickprov från fördelningar med väntevärden μ_1 respektive μ_2 och standardavvikelserna σ_1 respektive σ_2.

Om σ_1 och σ_2 är kända så är

$$I_{\mu_1-\mu_2} = (\bar{x} - \bar{y} - \lambda_{\alpha/2} D,\ \bar{x} - \bar{y} + \lambda_{\alpha/2} D)$$

ett tvåsidigt konfidensintervall för $\mu_1 - \mu_2$ med den approximativa konfidensgraden $1-\alpha$; här är $D = \sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}$.

Om σ_1 och σ_2 är okända tar man istället

$$I_{\mu_1-\mu_2} = (\bar{x} - \bar{y} - \lambda_{\alpha/2} d,\ \bar{x} - \bar{y} + \lambda_{\alpha/2} d)$$

där

$$d = \sqrt{s_1^2/n_1 + s_2^2/n_2}$$

SATS 12.β: Låt x vara en observation av $X \in Bin(n,p)$ där n är relativt stort och p är okänt. En punktskattning av p är den relativa frekvensen $p^*_{obs} = \frac{x}{n}$. Då är även

$$I_p = (p^*_{obs} - \lambda_{\alpha/2} d,\ p^*_{obs} + \lambda_{\alpha/2} d) \quad \text{där } d = \sqrt{p^*_{obs}(1-p^*_{obs})/n}$$

ett tvåsidigt konfidensintervall för p med den approximativa konfidensgraden $1-\alpha$.

SATS 12.γ: Låt x_1 och x_2 vara observationer av $X_1 \in Bin(n_1, p_1)$ respektive $X_2 \in Bin(n_2, p_2)$ där n_1 och n_2 är relativt stora och p_1 och p_2 är okända. Då har vi att $(p_1)^*_{obs} = x_1/n_1$, $(p_2)^*_{obs} = x_2/n_2$ och

$$I_{p_1-p_2} = ((p_1)^*_{obs} - (p_2)^*_{obs} - \lambda_{\alpha/2} d,\ (p_1)^*_{obs} - (p_2)^*_{obs} + \lambda_{\alpha/2} d)$$

där

$$d = \sqrt{\frac{(p_1)^*_{obs}(1-(p_1)^*_{obs})}{n_1} + \frac{(p_2)^*_{obs}(1-(p_2)^*_{obs})}{n_2}}$$

är ett tvåsidigt konfidensintervall för $p_1 - p_2$ med den approximativa konfidensgraden $1-\alpha$.

SATS 12.δ: SATS 12.β kan användas även om x är en observation a $X \in Hyp(N,n,p)$. Men då är $d = d_n\sqrt{p^*_{obs}(1-p^*_{obs})/n}$ där $d_n = \sqrt{(N-n)/(N-1)}$ eller $d_n \approx 1$ då $\frac{n}{N} < 0.1$.

SATS 12.ε: Låt x vara en observation av $X \in Po(\mu)$ där μ är okänt. Då är

$$I_\mu = (x - \lambda_{\alpha/2}\sqrt{x},\ x + \lambda_{\alpha/2}\sqrt{x})$$

ett tvåsidigt konfidensintervall för μ med den approximativa konfidensgraden $1-\alpha$.

SATS 12.ζ: Låt $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ där X_i är oberoende $N(\mu, \sigma)$ då är $\frac{(n-1)S^2}{\sigma^2} \in \chi^2(n-1)$.

SATS 12.η: Låt $X \in \chi^2(f)$, då är $P(\chi^2_{1-\alpha/2}(f) \leq X \leq \chi^2_{\alpha/2}(f)) = 1-\alpha$

SATS 12.θ: Låt $X_1, X_2, \dots, X_n \in N(\mu, \sigma)$, $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ och $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$ då är

$$\frac{\overline{X} - \mu}{S/\sqrt{n}} \in t(n-1)$$

Def(nollhypotes): Låt $\vec{x} = (x_1, \dots, x_n)$ vara ett slumpmässigt stickprov från en fördelning. Anta att fördelningen beror på en okänd parameter θ. Nollhypotesen H_0 innebär att man anger ett visst värde på θ, eventuellt flera värden. Om H_0 omfattar ett enda värde på θ kallas hypotesen enkel, annars sammansatt.

Def(testvariabel, kritiskt område, signifikanstest, signifikansnivå/felrisk): För att pröva H_0 introducerar man en testvariabel eller teststorhet $t_{obs} = t(\vec{x})$. Testvariabeln t_{obs} är en observation av en stickprovsvariabel $t(\vec{X})$. Vi anger också ett kritiskt område C, vilket är ett delområde av det område $t(\vec{X})$ kan anta värden i. Sedan används följande:

Signifikanstest: Om $t_{obs} \in C$ förkasta H_0
$t_{obs} \notin C$ förkasta ej H_0

Vi avpassar C så att

$$P(t(\vec{X}) \in C) = \alpha \quad \text{om } H_0 \text{ är sann}$$

där α väljs på förhand och kallas testets signifikansnivå, eller felrisk, och anger sannolikheten att H_0 förkastas om den är sann.

Def(signifikant, icke-signifikant): Om, efter ett signifikanstest med felrisk α, H_0 förkastas sägs resultatet vara signifikant på nivån α. Om H_0 ej förkastas sägs resultatet vara icke-signifikant på nivån α.
Man kan t.ex. säga att det finns en signifikant (eller ickesignifikant) avvikelse från nollhypotesen H_0 på nivån α.
O.B.S. Att H_0 ej förkastas betyder inte att den är sann utan endast att det ej kan uteslutas.

Def(signifikant*): Om man arbetar med signifikansnivåerna 0.05, 0.01 och 0.001 kan man använda kodbeteckningarna signifikant*, signifikant** respektive signifikant*** för att markera att ett resultat är signifikant på ifrågavarande nivå men inte för ett lägre α-värde bland de tre.

Def(P-värde/observerad signifikansnivå): Låt H_0 vara en nollhypotes som förkastas om testvariabeln $t = t_{obs} = t(\vec{x})$ är "tillräckligt stor" och stickprovsvariabeln $t(\vec{X})$ har en given känd sannolikhetsfördelning under H_0. P-värdet eller observerade signifikansnivån är sannolikheten $P = P(t(\vec{X}) \geq t(\vec{x}))$, beräknad under förutsättningen att H_0 är sann. Om P är "tillräckligt litet" förkastas H_0 då testresultatet ses som "osannolikt" då H_0 är sann.

SATS 13.α (Konfidensmetoden): Låt I_θ vara ett konfidensintervall för θ med konfidensgraden $1-\alpha$. Låt nollhypotesen H_0 vara att θ har det hypotetiska värdet θ_0. Om $\theta_0 \notin I_\theta$ är resultatet signifikant på nivån α, annars icke-signifikant.

Def(ensidigt, tvåsidigt): Om det kritiska området C har formen $[a,\infty)$ eller $(-\infty,b]$ säges signifikanstestet vara ensidigt, om C har formen $(-\infty,b]\cup[a,\infty)$ säges signifikanstestet vara tvåsidigt. $(a>b)$

Def(alternativ hypotes): En alternativ hypotes H_1, ibland kallad mothypotes, är en hypotes som kan vara sann om H_0 ej är sann. Med ett bra test menar man ett som förkastar H_0 med stor sannolikhet om H_1 skulle vara sann.

Def(styrkefunktion): Styrkefunktionen $h(\theta)$ definieras som
$h(\theta) = P(H_0 \text{ förkastas})$ om θ är det rätta parametervärdet
Ett test är bra om $h(\theta)$ är stor för alla $\theta \in H_1$ och liten för $\theta \in H_0$. Värdet $h(\theta)$ kallas testets styrka för ifrågavarande θ.
För ett och endast ett värde av θ i H_0, $\theta=\theta_0$, kommer testets konfidensgrad vara $\alpha = h(\theta_0)$. För flera värden av θ är konfidensgraden av testet $\alpha = \max_{\theta\in H_0} h(\theta)$

Def(tvåhypotesfall): Ett tvåhypotesfall är ett fall där antingen H_0 eller H_1 måste vara sann. I detta fall godkännes H_0 om H_0 ej förkastas med konfidensgrad α.

Def(flerhypotesfall): Ett flerhypotesfall är ett fall där vare sig H_0 eller H_1 behöver vara sann utan det kan finnas en (eller flera) övriga hypoteser H_k.

SATS 13.β: Låt $x_1,\dots,x_n$ vara ett slumpmässigt stickprov från $N(\mu,\sigma)$ och vi vill pröva den enkla hypotesen $H_0: \mu=\mu_0$.
Vi tar då testvariabel:
$$u = \begin{cases} (\bar{x}-\mu_0)/D & \text{om } \sigma \text{ är känt, } D=\sigma/\sqrt{n} \\ (\bar{x}-\mu_0)/d & \text{om } \sigma \text{ är okänt, } d=s/\sqrt{n} \end{cases}$$
Om H_0 är sann är $U\in N(0,1)$ och $U\in t(n-1)$ om σ är känd respektive okänd.
Om σ är känd utför följande test beroende på formen av H_1:
- a) Om H_1 omfattar värden på μ som både är större och mindre än μ_0 används ett tvåsidigt test av formen: Om $|u| \geq \lambda_{\alpha/2}$ förkasta H_0 med signifikansnivå α.
- b) Om H_1 omfattar värden på μ som endast är större än μ_0 används ett ensidigt test av formen: Om $u \geq \lambda_\alpha$ förkasta H_0 med signifikansnivå α.
- c) Om H_1 omfattar värden på μ som endast är mindre än μ_0 används ett ensidigt test av formen: Om $u \leq -\lambda_\alpha$ förkasta H_0 med signifikansnivå α.

Om σ är okänd utför istället följande test:
- a) Om H_1 omfattar värden på μ som både är större och mindre än μ_0 används ett tvåsidigt test av formen: Om $|u| \geq t_{\alpha/2}(n-1)$ förkasta H_0 med signifikansnivå α.
- b) Om H_1 omfattar värden på μ som endast är större än μ_0 används ett ensidigt test av formen: Om $u \geq t_\alpha(n-1)$ förkasta H_0 med signifikansnivå α.
- c) Om H_1 omfattar värden på μ som endast är mindre än μ_0 används ett ensidigt test av formen: Om $u \leq -t_\alpha(n-1)$ förkasta H_0 med signifikansnivå α.

SATS 13.γ: Låt $x_1,\dots,x_n$ vara ett slumpässigt stickprov från en fördelning. Ta en testvariabeln av form
$$u = \begin{cases} (\theta^*_{obs}-\theta_0)/D & \text{om } D \text{ är helt känt} \\ (\theta^*_{obs}-\theta_0)/d & \text{om } D \text{ ej är helt okänt} \end{cases}$$
Om U kan antas vara ungefär normalfördelad kan samma test användas som för ett känt σ i SATS 13.β, dock kan inte signifikansnivån α bestämmas exakt med denna metod.

SATS 13.δ (χ^2-test av given fördelning): Låt n oberoende försök göras där vart och ett av försöken kan utfalla på r olika sätt $A_1, A_2, \dots, A_r$ med sannolikheter $P(A_1), P(A_2), \dots, P(A_r)$ där $\sum P(A_j) = 1$. Låt vidare $x_1, x_2, \dots, x_r$ ($\sum x_j = n$) vara de absoluta frekvenserna för alternativen $A_1, A_2, \dots, A_r$. Låt H_0 vara
$H_0: P(A_1) = p_1,\ P(A_2) = p_2, \dots, P(A_r) = p_r$ där $p_1, p_2, \dots, p_r$ är givna värden så att $\sum p_j = 1$.
Låt vidare

$$Q_{obs} = \sum_{j=1}^{r} \frac{(x_j - np_j)^2}{np_j}$$

Om H_0 är sann och $np_j \geq 5\ \forall j$ så är Q_{obs} en observation från en fördelning som är approximativt $\chi^2(r-1)$.
Om $np_j \geq 5\ \forall j$ har vi signifikanstestet: Förkasta H_0 om $Q_{obs} > \chi^2_\alpha(r-1)$
Förkasta ej H_0 om $Q_{obs} \leq \chi^2_\alpha(r-1)$
Detta test har en approximativ signifikansnivå α.

SATS 13.ε (χ^2-test av fördelning med skattade parametrar): Låt oss ha samma scenario som i SATS 13.δ med skillnaden att fördelningen är okänd och det är angående fördelningen vi har en hypotes, utan att bestämma
$H_0: P(A_1) = p_1(\theta), \dots, P(A_r) = p_r(\theta)$ där $\theta = (\theta_1, \dots, \theta_k)$ är k okända parametrar
Om man skattar θ till $\theta^*_{obs} = ((\theta_1)^*_{obs}, \dots, (\theta_k)^*_{obs})$ och låter $p_j^* = p_j(\theta^*_{obs})$ är

$$Q_{obs} = \sum_{j=1}^{r} \frac{(x_j - np_j^*)^2}{np_j^*}$$

en observation av en fördelning som är approximativt $\chi^2(r-k-1)$

om $np_j^* \geq 5\ \forall j$ och H_0 är sann.
Om $np_j^* \geq 5\ \forall j$ har vi signifikanstestet: Förkasta H_0 om $Q_{obs} > \chi^2_\alpha(r-k-1)$
Förkasta ej H_0 om $Q_{obs} \leq \chi^2_\alpha(r-k-1)$
Detta test har en approximativ signifikansnivå α.

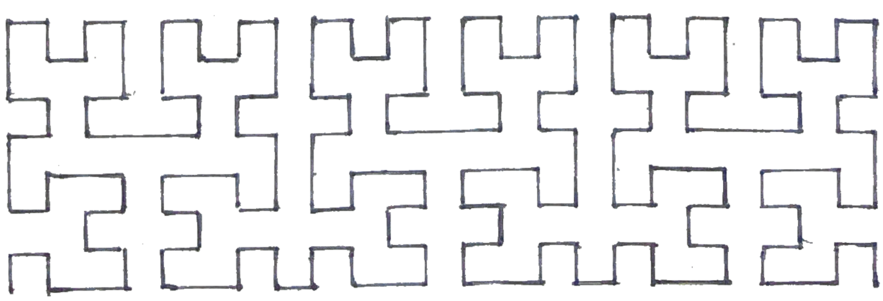
SATS 13.ζ (χ^2-Homogenitetstest): Anta att man utfört s serier av oberoende försök och vill undersöka om serierna kan anses vara homogena, alltså att samma uppsättning sannolikheter $P(A_1), \dots, P(A_r)$ förekommer i alla serier.
Antag att s serier utfördes med följande resultat

Serie	Absoluta frekvenser för A_1	A_2	$\dots$	A_r	Antal försök
1	x_{11}	x_{12}	$\dots$	x_{1r}	n_1
2	x_{21}	x_{22}	$\dots$	x_{2r}	n_2
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
s	x_{s1}	x_{s2}	$\dots$	x_{sr}	n_s
Summa	$x_{\cdot 1}$	$x_{\cdot 2}$	$\dots$	$x_{\cdot r}$	n

Bilda testvariabeln

$$Q_{obs} \sum_{i=1}^{s} \sum_{j=1}^{r} \frac{(x_{ij} - n_i p_j^*)^2}{n_i p_j^*}$$

där $p_j^* = x_{\cdot j}/n$.
Om H_0: {serierna är homogena} har vi homogenitetstestet:
Förkasta hypotesen om homogenitet om $Q_{obs} > \chi^2_\alpha((r-1)(s-1))$
Om $n_i p_j^* \geq 5\ \forall i,j$ har testet en approximativ signifikansnivå α.



SATS 13.η (χ^2-Oberoendetest): Vi vill testa om två egenskaper A och B är oberoende. Vi utför N oberoende försök som kan utfalla på sr olika sätt B_iA_j för $i=1,\dots,s$ och $j=1,\dots,r$ med sannolikheten respektive p_{ij} ($\sum_i \sum_j p_{ij}=1$). Låt x_{ij} vara absoluta frekvensen för B_iA_j ($\sum_i \sum_j x_{ij}=N$).
Resultatet av försöken presenteras i en kontingenstabell:

Antal observationer	A_1	A_2	$\dots$	A_r	Radsumma
B_1	x_{11}	x_{12}	$\dots$	x_{1r}	n_1
B_2	x_{21}	x_{22}	$\dots$	x_{2r}	n_2
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
B_s	x_{s1}	x_{s2}	$\dots$	x_{sr}	n_s
Kolonnsumma	m_1	m_2	$\dots$	m_r	N

Låt $Q_{obs}=\sum_{i=1}^{s}\sum_{j=1}^{r}\dfrac{\left(x_{ij}-\frac{n_im_j}{N}\right)^2}{\frac{n_im_j}{N}}$.

Låt H_0: {A och B är oberoende}
Vi kan då använda oberoendetestet:
Förkasta hypotesen om oberoende om $Q_{obs}>\chi^2_\alpha((r-1)(s-1))$
Om $\frac{n_im_j}{N}\geq 5$ $\forall i,j$ har testet en approximativ signifikansnivå α.

Def(enkel linjär regression, teoretisk regressionslinje, regressionsvariabel): Observera följande modell vilket är ett exempel av linjär regression: Det föreligger n par av värden $(x_1,y_1),\dots,(x_n,y_n)$ där $x_1,\dots,x_n$ är givna storheter och $y_1,\dots,y_n$ är observationer av oberoende s.v. $Y_1,\dots,Y_n$ där $Y_i\in N(\mu_i,\sigma)$. Varje väntevärde μ_i är linjärt beroende av x_i, dvs. $\mu_i=\alpha+\beta x_i$ $i=1,\dots,n$.
Linjen $y=\alpha+\beta x$ kallas den teoretiska regressionslinjen. Den visar hur väntevärdet beror av regressionsvariabeln x.

SATS 14.α: Med hjälp av MK-metoden fås följande skattningar av α och β:

$$\beta^*=\frac{S_{xy}}{S_{xx}} \quad \text{och} \quad \alpha^*=\bar{y}-\beta^*\bar{x}$$

där

$$S_{xy}=\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y}) \quad \text{och} \quad S_{xx}=\sum_{i=1}^{n}(x_i-\bar{x})^2$$

Def(skattad regressionslinje): Genom att sätta in α^* och β^* från SATS 14.α i den teoretiska regressionslinjen får man den skattade regressionslinjen
$y=\alpha^*+\beta^*x$
För varje givet $x=x_0$ kan man med denna linje skatta det tillhörande väntevärdet $\mu_0=\alpha+\beta x_0$. Vi kallar denna skattning μ_0^* och har alltså $\mu_0^*=\alpha^*+\beta^*x_0$.
Denna procedur kallas skattning av punkt på den teoretiska regressionslinjen eller, kortare, skattning av väntevärde.

SATS 14.β: Räkneregler för väntevärden och varianser ger oss:
$E(\beta^*)=\beta$ $\quad V(\beta^*)=\sigma^2/S_{xx}$ $\quad E(\mu_0^*)=\mu_0$ $\quad V(\mu_0^*)=\sigma^2\left(\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right)$

Def(residualkvadratsumma): Vi har $Q(\alpha,\beta)=\sum_{i=1}^{n}(y_i-\mu_i)^2=\sum_{i=1}^{n}(y_i-\alpha-\beta x_i)^2$ som har minimivärdet $Q_0=\sum_{i=1}^{n}(y_i-\alpha^*-\beta^*x_i)^2$ som kallas residualkvadratsumman.

SATS 14.γ: Vi har att $Q_0=S_{yy}-\beta^{*2}S_{xx}=S_{yy}-S_{xy}^2/S_{xx}$ där $S_{yy}=\sum_{i=1}^{n}(y_i-\bar{y})^2$.
En skattning av σ (från definitionen för enkel linjär regression) är den korrigerade ML-skattningen: $s=\sqrt{Q_0/(n-2)}$. σ^2 kan skattas som $s=Q_0/(n-2)$.

SATS 14.5: Om σ är känt har vi konfidensintervallen

$I_\beta = (\beta^* - \lambda_{p/2} D,\ \beta^* + \lambda_{p/2} D);\ (D = \sigma / \sqrt{\Sigma (x_i - \bar{x})^2})$

och

$I_{\mu_0} = (\mu_0^* - \lambda_{p/2} D,\ \mu_0^* + \lambda_{p/2} D);\ \left(D = \sigma \Big/ \sqrt{\frac{1}{n} + \frac{(x_0 - \bar{x})^2}{S_{xx}}}\right)$

för β respektive μ_0 med konfidensgraden $1-p$

Om σ är okänt har vi istället konfidensintervallen (med samma konfidensgrad):

$I_\beta = (\beta^* - t_{p/2}(n-2)d,\ \beta^* + t_{p/2}(n-2)d);\ (d = s/\sqrt{S_{xx}})$

och

$I_{\mu_0} = (\mu_0^* - t_{p/2}(n-2)d,\ \mu_0^* + t_{p/2}(n-2)d);\ (d = s\sqrt{1/n + (x_0 - \bar{x}^2)/S_{xx}}$

Där

$s = \sqrt{Q_0/(n-2)}$