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 reelvärd 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

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n$$ 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 $0 \leq k \leq v$, $0 \leq n-k \leq s$ 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

$$p_X(k) = \frac{\mu^k}{k!} e^{-\mu}, \quad k = 0, 1, 2, \ldots, \quad \mu > 0$$ 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
$$f_X(x) \begin{cases} \lambda e^{-\lambda x} & \text{om } x > 0 \\ 0 & \text{om } x \leq 0 \end{cases}$$
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
$$f_X(x) = \begin{cases} \lambda c(\lambda x)^{c-1} e^{-(\lambda x)^c} & x > 0 \\ 0 & x \leq 0 \end{cases}$$
där λ och c är positiva tal, säges X vara Weibull-fördelad

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

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
$$f_X(x) = \begin{cases} \frac{\lambda^c}{\Gamma(c)} \cdot x^{c-1} e^{-\lambda x} & \text{om } x > 0 \\ 0 & \text{om } x \leq 0 \end{cases}$$
där $c > 0$ och $\lambda > 0$, säges X vara gammafördelad

Def(fördelningsfunktion): $F_X(x) = P(X \leq x)$, $x \in \mathbb{R}$, 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 $A_1, A_2, \ldots, A_r$ med sannolikheter $p_1, \ldots, p_r$ där $p_1 + \ldots + p_r = 1$. Låt X_j vara antalet gånger A_j inträffar vid n oberoende upprepningar av försöktet. Då gäller att $X_j \in Bin(n, p_j)$. Då har vi att den simultana sannolikhetsfunktionen för den r-dimensionella s.v. $(X_1, \ldots, X_r)$ ges av

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

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

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

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

$$F_{X,Y}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u,v)\,du\,dv$$ om A 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 $Z_1 = \max(X_1, \ldots, X_n)$ och $Z_2 = \min(X_1, \ldots, X_n)$ där $X_1, \ldots, X_n$ 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 $V(X)=E((X-\mu)^2)$
Alltså
$$V(X)=\begin{cases}\sum_k(k-\mu)^2p_X(k) & \text{då } X \text{ är en diskret s.v.}\\ \int_{-\infty}^{\infty}(x-\mu)^2f_X(x)\,dx & \text{då } X \text{ är en kontinuerlig s.v.}\end{cases}$$

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

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:
$C(X,Y) = E(XY) - E(X)E(Y) = E(XY) - \mu_X\mu_Y$ 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. $X_1, \dots, X_n$ 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. $X_1, \dots, X_n$ 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 $X_1, \dots, X_n$ är en s.v. med samma väntevärde μ så gäller att

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

Och om $X_1, \dots, X_n$ 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 $X_1, \dots, X_n$ 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 $X_1, X_2, \dots$ 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
$$P\left(a < \frac{Y_n - n\mu}{\sigma\sqrt{n}} \le b\right) \to \Phi(b) - \Phi(a) \text{ då } n \to \infty.$$

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
$$P\left(a < \frac{Z_n - A_n}{B_n} \le b\right) \to \Phi(b) - \Phi(a) \text{ då } n \to \infty$$
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 $P(a < X \le b)$ approximeras som
$$P(a < X \le b) \approx \Phi\left(\frac{b - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a - np}{\sqrt{np(1-p)}}\right) \quad \text{för } np(1-p) \ge 10 \text{ och } a, b \in \mathbb{N}$$
eller, för mer exakta värden:
$$P(a < X \le b) \approx \Phi\left(\frac{b + 1/2 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{a + 1/2 - np}{\sqrt{np(1-p)}}\right) \quad \text{för } np(1-p) \ge 3 \text{ och } a, b \in \mathbb{N}$$
Alltså
- $Bin(n,p) \sim N(np, \sqrt{np(1-p)}) \quad : np(1-p) \ge 10$
- $Bin(n,p) \sim N(np - 1/2, \sqrt{np(1-p)}) : np(1-p) \ge 3$

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

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 $0 \le Y \le 1$ och
$E(Y) = p, \quad V(Y) = p(1-p)/n, \quad D(Y) = \sqrt{p(1-p)/n}$

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 $0 \le k \le Np$, $0 \le n-k \le N(1-p)$

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 $\frac{N-n}{N-1} \cdot np(1-p) \ge 10$
- $X \in Po(np)$ 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 tidsenhet, 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 $X \in N(\mu, \sqrt{\mu})$