MS-A0509 Grundkurs i sannolikhetskalkyl och statistik
Sammanfattning, del I

G. Gripenberg

Aalto-universitetet

6 februari 2015

Vad är sannolikhet?

- *Relativ frekvens vid upprepningar: Om en fabrik tillverkat* 1000 000 *exemplar av en produkt av vilka* 5015 *har något fel så är sannolikheten för en felaktighet* 0.005
- *Andelen fall då ett något förekommer: Om i en urna finns* 6 *svarta och* 4 *vita kulor och man slumpmässigt väljer en kula så är sannolikheten att den är svart* $\frac{6}{6+4} = 0.6$.
- *Ett mått på hur troligt man anser något vara: "Sannolikheten för hård vind imorgon är* 70%*."*

Sannolikhet, händelser, utfallsrum

- *Mängden av alla tänkbara resultat av ett "experiment" eller ett "slumpmässigt försök" är* **utfallsrummet**, *ofta betecknat med* Ω.
- *Elementen i utfallsrummet, dvs. enskilda resultat av experimentet är* **elementarhändelser**.
- **Händelser** *är delmängder av utfallsrummet och när man säger att händelsen A inträffar, menar man alltid att någon elementarhändelse som hör till A inträffar.*
- *För varje händelse $A \subset \Omega$ finns det en sannolikhet* $\Pr(A)$.
- *Sannolikhetsfunktionen skall uppfylla följande villkor:*
  - ★ $0 \leq \Pr(A) \leq 1$ *för varje händelse A.*
  - ★ $\Pr(\Omega) = 1$.
  - ★ $\Pr(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \Pr(A_i)$ *om $A_j \cap A_k = \emptyset$ då $j \neq k$.*

  Då gäller också följande:
  - ★ $\Pr(\emptyset) = 0$.
  - ★ $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$.
  - ★ $\Pr(\Omega \setminus A) = 1 - \Pr(A)$.
  - ★ $A \subset B \Rightarrow \Pr(A) \leq \Pr(B)$.

Oberoende

Händelserna A och B är oberoende ifall

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B),$$

och händelserna A_j, $j \in J$ är oberoende om

$$\Pr(A_{j_1} \cap A_{j_2} \cap \ldots \cap A_{j_m}) = \Pr(A_{j_1}) \cdot \Pr(A_{j_2}) \cdot \ldots \cdot \Pr(A_{j_m})$$

alltid då $j_k \in J$, $k = 1, \ldots, m$, $j_p \neq j_q$ då $p \neq q$.

Obs!

Om händelserna A_j, $j \in J$ är oberoende så är A_{j_p} och A_{j_q} oberoende då $j_p \neq j_k$ men om A_{j_p} och A_{j_q} är oberoende för alla $j_p \neq j_k$ så behöver inte händelserna A_j, $j \in J$ vara oberoende.

Betingad sannolikhet

Den betingade sannolikheten för händelsen A givet händelsen B är

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

då man antar att $\Pr(B) > 0$.

Då händelsen B är given kan man begränsa utfallsrummet från Ω till B och räkna om sannolikheterna för händelserna $A \cap B$ som är delmängder av det nya utfallsrummet.

Produktregeln för betingad sannolikhet

Av definitionen för betingad sannolikhet följer den sk. produktregeln

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B|A),$$

och mera allmänt

$$\Pr(A_1 \cap \ldots \cap A_k) = \Pr(A_1) \cdot \Pr(A_2|A_1) \cdot \Pr(A_3|A_1 \cap A_2) \cdot \ldots \cdot \Pr(A_k|A_1 \cap \ldots \cdot \cap A_{k-1}).$$

Total sannolikhet

Om $\cup_{j=1}^n A_j = \Omega$, $A_j \cap A_k = \emptyset$ *då* $j \neq k$ *och* $\Pr(A_j) > 0$ *då* $j = 1, \ldots, n$ *så gäller*

$$\Pr(B) = \sum_{j=1}^{n} \Pr(A_j) \cdot \Pr(B|A_j).$$

Varför? Eftersom $B = B \cap \Omega = \cup_{j=1}^n B \cap A_j$ *och* $(B \cap A_j) \cap (B \cap A_k) = \emptyset$ *då* $j \neq k$ *så är* $\Pr(B) = \sum_{j=1}^n \Pr(B \cap A_j)$ *och enligt definitionen är* $\Pr(A_j) \cdot \Pr(B|A_j) = \Pr(B \cap A_j)$.

Bayes formel

Om $\cup_{j=1}^n A_j = \Omega$, $A_j \cap A_k = \emptyset$ *då* $j \neq k$, $\Pr(B) > 0$ *och* $P(A_j) > 0$, $j = 1, \ldots, n$ *så gäller*

$$\Pr(A_k|B) = \frac{\Pr(A_k) \cdot \Pr(B|A_k)}{\sum_{j=1}^n \Pr(A_j) \cdot \Pr(B|A_j)}.$$

Varför?

$$\Pr(A_k|B) = \frac{\Pr(A_k \cap B)}{\Pr(B)}, \quad \Pr(A_k) \cdot \Pr(B|A_k) = \Pr(A_k \cap B) \quad \textit{och}$$

$$\sum_{j=1}^{n} \Pr(A_j) \cdot \Pr(B|A_j) = \Pr(B).$$

Klassisk sannolikhet och kombinatorik

$$\Pr(A) = \frac{\textit{Antal fall då A inträffar}}{\textit{Totala antalet möjliga fall}}$$

Man antar alltså att varje elementarhändelse är lika sannolik och problemet blir att bestämma hur många element det det finns i utfallsrummet Ω och hur många av dessa hör som till mängden A.

Produktprincipen

Om i en urvalsprocess finns k steg och i steg j finns n_j alternativ, oberoende av vilka val som gjorts i tidigare steg (men vilka alternativen är kan bero på valen) så är det totala antalet alternativ

$$n_1 \cdot n_2 \cdot \ldots \cdot n_k.$$

Permutationer, binomialkoefficienter etc.

- *Om det i en mängd finns n element kan dessa ordnas på*

$$n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (n-1) \cdot n,$$

olika sätt. (Kom ihåg: $0! = 1$*)*

- *Om man ur en mängd med n element väljer k element och beaktar i vilken ordning elementen väljs, kan detta göras på*

$$n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}$$

olika sätt.

- *Om man ur en en mängd med n element väljer en delmängd med k element, dvs. inte beaktar i vilken ordning elementen väljs, kan detta göras på*

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

olika sätt. Av dethär följer att om ett experiment upprepas n gånger så att händelser vid olika gånger är oberoende så är sannolikheten för att händelsen A inträffar exakt k gånger $\binom{n}{k} \Pr(A)^k \big(1 - \Pr(A)\big)^{n-k}$.

Plocka kulor med eller utan återläggning

Antag att i en urna finns s svarta och v vita kulor och att vi plockar n kulor ur urnan.

(a) *Om vi för varje kula noterar vilken färg den har och sedan lägger den tillbaka i urnan så använder vi återläggning. Sannolikheten att vi plockar en svart kula är $\frac{s}{s+v}$ och för en vit är den $\frac{v}{s+v}$ så att sannolikheten att vi plockar k svarta och $n-k$ vita i en viss given ordning är $\left(\frac{s}{s+v}\right)^k \left(\frac{v}{s+v}\right)^{n-k}$ och då är sannolikheten att vi plockar k svarta och $n-k$ vita i vilken ordning som helst*

$$\binom{n}{k}\left(\frac{s}{s+v}\right)^k \left(\frac{v}{s+v}\right)^{n-k}.$$

(b) *Om vi däremot inte använder återläggning så kommer sannolikheten att vi plockar en svart kula att bero på vilka kulor vi redan plockat och sannolikheten att vi plockar k svarta och $n-k$ vita är $\frac{\binom{s}{k}\cdot\binom{v}{n-k}}{\binom{s+v}{n}}$ eftersom vi kan plocka k svarta bland s svarta på $\binom{s}{k}$ olika sätt och $n-k$ vita bland v vita på $\binom{v}{n-k}$ olika sätt.*

Slumpvariabler och fördelningsfunktioner

En (reell) **slumpvariabel** *(eller* **stokastisk variabel***) är en funktion $X : \Omega \to \mathbb{R}$ (alltså inte egentligen en variabel) där Ω är ett utfallsrum för ett experiment i vilken en sannolikhet är definierad och $\{\, \omega \in \Omega : X(\omega) \le t \,\}$ är en händelse för alla $t \in \mathbb{R}$.*

Om X är en (reell) slumpvariabel så är dess (kumulativa) fördelningsfunktion funktionen

$$F_X(t) = \Pr(X \le t) = \Pr\big(\{\, \omega \in \Omega : X(\omega) \le t \,\}\big).$$

En funktion $F : \mathbb{R} \to [0,1]$ är en fördelningsfunktion om och endast om

- $0 \le F(s) \le F(t) \le 1$ *då* $s < t$,
- $\lim_{t\to-\infty} F(t) = 0$ *och* $\lim_{t\to\infty} F(t) = 1$,
- $\lim_{s\to t+} F(s) = F(t)$ *då* $t \in \mathbb{R}$.

När F är en fördelningsfunktion för X så gäller dessutom att

- $F(t) - F(s) = \Pr(s < X \le t)$ *då* $s < t$,
- $\lim_{s\to t-} F(s) = \Pr(X < t)$,
- $\lim_{s\to t+} F(s) - \lim_{s\to t-} F(s) = F(t) - \lim_{s\to t-} F(s) = \Pr(X = t)$.

Obs!

Uttryck som $X \leq t$ *och* $X < t$ *är formellt sett inte händelser (dvs. delmängder i* Ω*) men man skriver oftast* $\Pr(X \leq t)$ *istället för det längre uttrycket* $\Pr\big(\{\,\omega \in \Omega : X(\omega) \leq t\,\}\big)$.

Oberoende slumpvariabler

De (reella) slumpvariablerna X_j, $j \in J$ *definierade i samma utfallsrum är oberoende om händelserna* $\{\,X_j \leq a_j\,\}$, $j \in J$ *är oberoende för alla* $a_j \in \mathbb{R}$, $j \in J$ *och då är också händelserna* $\{\,X_j \in A_j\,\}$, $j \in J$ *oberoende för alla Borel mängder* A_j.

En slumpvariabels sannolikhetsfördelning

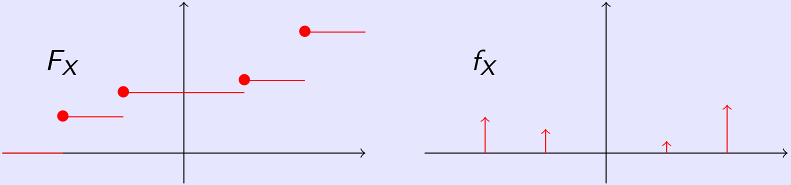
Slumpvariabelns $X : \Omega \to \mathbb{R}$ *sannolikhetsfördelning (eller bara fördelning) är sannolikhetsfunktionen* $\Pr_X(A) = \Pr(X \in A)$ *där* $A \subset \mathbb{R}$ *är sådan att* $\{\,\omega \in \Omega : X(\omega) \in A\,\}$ *är en händelse dvs. mängden reella tal är utfallsrummet och sannolikheterna för dess händelser definieras med funktionen* $\Pr_X$. *Om slumpvariabelns* X *fördelning är tex. den sk. normalfördelningen med parameterna* μ *och* σ^2 *så skriver man dess fördelningsfunktion som* $F_{\mathrm{N}(\mu,\sigma^2)}$ *istället för* F_X.

Diskreta slumpvariabler

En (reell) slumpvariabel X är diskret om det finns en mängd $A \subset \mathbb{R}$ och positiva tal $f_X(a)$, $a \in A$ så att

$$F_X(t) = \sum_{\substack{a \le t \\ a \in A}} f_X(a).$$

Detta innebär att $\Pr(X = a) = f_X(a)$ *då* $a \in A$ *och* $\sum_{a \in A} f_X(a) = 1$ *så att* $\Pr(X \notin A) = 0$ *och mängden A innehåller högst numrerbart många element och vi kan anta att $f_X(t) = 0$ då $t \notin A$.*
Funktionen f_X är **frekvensfunktionen** *eller sannolikhetsfunktionen för X.*

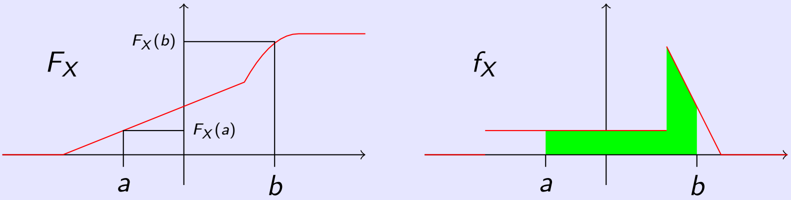


Kontinuerliga slumpvariabler

En slumpvariabel X är kontinuerlig om fördelningsfunktionen är kontinuerlig, dvs. om $\Pr(X = a) = 0$ för alla $a \in \mathbb{R}$. Oftast antar man ändå att slumpvariabeln X har en **täthetsfunktion** *f_X så att*

$$F_X(t) = \int_{-\infty}^{t} f_X(s)\,\mathrm{d}s.$$

Detta innebär att $f_X(s) \geq 0$ och $\int_{-\infty}^{\infty} f_X(s)\,\mathrm{d}s = 1$.



$$\Pr(a \leq X \leq b) = \Pr(a < X < b) = F_X(b) - F_X(a) = \int_{a}^{b} f_X(s)\,\mathrm{d}s.$$

Väntevärde

Om X är en diskret slumpvariabel med frekvensfunktion f_X så är dess väntevärde

$$E(X) = \sum_a a \Pr(X = a) = \sum_a a f_X(a),$$

och om X är en kontinuerlig slumvariabel med täthetsfunktion f_X så är dess väntevärde

$$E(X) = \int_{-\infty}^{\infty} s f_X(s)\, ds,$$

i båda fallen förutsatt att summan eller integralen existerar (dvs. $\sum_{a>0} a f_X(a) < \infty$ eller $\sum_{a<0} |a| f_X(a) < \infty$ och $\int_0^\infty t f_X(t)\, dt < \infty$ eller $\int_{-\infty}^0 |t| f(t)\, dt < \infty$), i annat fall skriver man $E(X) = NaN$ *och säger att slumpvariabeln inte har något väntevärde.*

Väntevärdet av en funktion av en slumpvariabel

Om X är en diskret slumpvariabel och g är en mätbar *funktion så är*

$$E(g(X)) = \sum_a g(a)\Pr(X = a) = \sum_a g(a) f_X(a)$$

och om X är en kontinuerlig slumpvariabel med täthetsfunktion f_X så är

$$E(g(X)) = \int_{-\infty}^{\infty} g(s) f_X(s)\,ds.$$

Om $g(t) = 1$ då $t \in A$ och $g(t) = 0$ annars (och då skriver man ofta $g = \mathbf{1}_A$) så är

$$E(g(X)) = \Pr(X \in A),$$

dvs. också sannolikheter kan skrivas som väntevärden.

Varians och standardavvikelse

Om slumpvariabeln X *har ett väntevärde så är dess* **varians**

$$\mathrm{Var}(X) = \sigma_X^2 = \mathrm{E}\Big(\big(X - \mathrm{E}(X)\big)^2\Big),$$

och dess **standardavvikelse** *är on*

$$\mathrm{D}(X) = \sigma_X = \sqrt{\mathrm{Var}(X)}.$$

(Observera att variansen aldrig är negativ!)
Fördelen med standardavvikelsen är att den har samma enhet som X *och att* $\mathrm{D}(\alpha X) = |\alpha|\mathrm{D}(X)$ *då* α *är något reellt tal medan* $\mathrm{Var}(\alpha X) = \alpha^2\mathrm{Var}(X)$ *och* $\mathrm{Var}(X)$ *har enheten* m^2 *om* X *tex. har enheten* m *(men variansen har andra stora fördelar).*

Väntevärdet är linjärt och monotont

Ifall X_1 och X_2 är två slumpvariabler (definierade i samma utfallsrum), som har ändliga väntevärden och c_1 och c_2 är reella tal så är

$$\mathrm{E}(c_1X_1 + c_2X_2) = c_1\mathrm{E}(X_1) + c_2\mathrm{E}(X_2),$$

och om dessutom $\Pr(X_1 \leq X_2) = 1$ *så är*

$$\mathrm{E}(X_1) \leq \mathrm{E}(X_2).$$

En följd av dethär är att (där $\mathbf{1}$ *är en slumpvariabel som får värdet* 1 *med sannolikheten* 1*)*

$$\begin{aligned}\mathrm{Var}(X) &= \mathrm{E}\big((X - \mathrm{E}(X))^2\big) = \mathrm{E}\big(X^2 - 2X\mathrm{E}(X) + \mathrm{E}(X)^2\big) \\ &= \mathrm{E}(X^2) - 2\mathrm{E}(X)\mathrm{E}(X) + \mathrm{E}(X)^2\mathrm{E}(\mathbf{1}) = \mathrm{E}(X^2) - \mathrm{E}(X)^2.\end{aligned}$$

Variansen av summan av två slumpvariabler

Ifall X_1 och X_2 är **oberoende** *slumpvariabler (definierade i samma utfallsrum) så är*

$$\mathrm{Var}(c_1X_1 + c_2X_2) = c_1^2\mathrm{Var}(X_1) + c_2^2\mathrm{Var}(X_2),$$

om c_1 och c_2 är reella tal och i allmänhet (förutsatt att varianserna är ändliga) gäller

$$\begin{aligned}\mathrm{Var}(c_1X_1 + c_2X_2) &= c_1^2\mathrm{Var}(X_1) \\ &\quad + 2c_1c_2\mathrm{E}\big((X_1 - \mathrm{E}(X_1))(X_2 - \mathrm{E}(X_2))\big) + c_2^2\mathrm{Var}(X_2).\end{aligned}$$

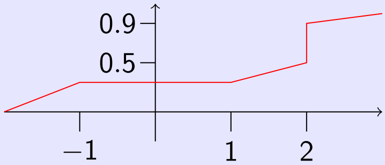
Kvantiler

Antag att X är en slumpvariabel med fördelningsfunktion F_X och $0 < p < 1$.

- *Om F_X har en invers funktion så är $x_p = F_X^{-1}(p)$ slumpvariabelns X och dess fördelnings p-***kvantil***.*
- *I allmänhet är x_p en p-***kvantil*** ifall*

$$\Pr(X < x_p) \leq p \leq \Pr(X \leq x_p),$$
$$\Pr(X > x_p) \leq 1 - p \leq \Pr(X \geq x_p).$$

- **Medianen** *är en* 0.5*-kvantil.*
- *Kvantilerna är inte nödvändigtvis entydiga men de existerar alltid.*
- *Ofta väljer man som p-kvantil mittpunkten på intervallet med alla p-kvantiler.*

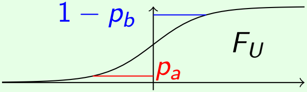


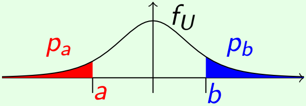
$$x_{0.4} \in [-1, 1]$$
$$x_q = 2, \quad q \in [0.5, 0.9]$$

Kvantiler, forts.

I många beräkningar i statistik bildar vi först en "testvariabel" U, vars fördelningsfunktion F_U och täthetsfunktion f_U vi känner till (åtminstone approximativt). Sedan bestämmer vi tal a och b så att $\Pr(U < a) = p_a$ *och* $\Pr(U > b) = p_b$ *där vanligtvis $p_a = p_b$ men ibland är det ena talet* 0*. Om fördelningsfunktionen F_U har en invers funktion så får vi*

$$a = F_U^{-1}(p_a) \quad \textit{och} \quad b = F_U^{-1}(1 - p_b)$$





Med hjälp av dessa tal a och b och definitionen av testvariabeln kan vi sedan räkna ut det vi verkligen är intresserade av.
Ifall U som här är kontinuerlig så är $\Pr(U < a) = \Pr(U \leq a)$ *och* $\Pr(a \leq U \leq b) = \Pr(a < U < b) = 1 - p_a - p_b$.

Några viktiga diskreta slumpvariabler och deras fördelningar

- *Jämn diskret fördelning:* $\Pr(X = x_i) = \frac{1}{n}$, $i = 1, 2, \ldots, n$ *och man antar att $x_i \neq x_j$ då $i \neq j$.*
- *Bernoullifördelning* $X \sim \text{Bernoulli}(p)$, $0 \leq p \leq 1$:
  - ◇ $\Pr(X = 1) = p$, $\Pr(X = 0) = (1 - p)$ *dvs. $X(\omega) = 1$ då $\omega \in A \subset \Omega$ och $X(\omega) = 0$ då $\omega \in \Omega \setminus A$ där $\Pr(A) = p$.*
  - ◇ $\mathrm{E}(X) = p$ *och* $\mathrm{Var}(X) = p(1 - p)$.
- *Binomialfördelning* $X \sim \text{Bin}(n, p)$, $n \geq 0$, $0 \leq p \leq 1$:
  - ◇ $\Pr(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$, $k = 0, 1, \ldots, n$.
  - ◇ *X är summan av n oberoende* Bernoulli*(p)-fördelade slumvariabler, dvs. experimentet upprepas n gånger med oberoende resultat och händelsen A, med sannolikheten p inträffar X gånger.*
  - ◇ $\mathrm{E}(X) = np$ *och* $\mathrm{Var}(X) = np(1 - p)$.
- *Poisson-fördelning* $X \sim \text{Poisson}(\lambda)$, $\lambda \geq 0$:
  - ◇ $\Pr(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$, $k = 0, 1, 2, \ldots$.
  - ◇ *Fås som gränsvärde av binomialfördelningen då $n \to \infty$ och $np \to \lambda$.*
  - ◇ $\mathrm{E}(X) = \mathrm{Var}(X) = \lambda$.

Några viktiga diskreta slumpvariabler och deras fördelningar, forts.

- *Hypergeometrisk fördelning* $X \sim \text{HyperGeom}(N, r, n)$ $0 \leq n, r \leq N$*:*
  - $\Pr(X = k) = \dfrac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}$.
  - *Om man plockar n kulor ur en urna som innehåller r vita och $N - r$ svarta kulor så är X antalet vita kulor man plockat.*
  - $E(X) = \frac{nr}{N}$, $\text{Var}(X) = \frac{nr}{N} \cdot \frac{N-r}{N} \cdot \frac{N-n}{N-1}$.
- *Geometrisk fördelning* $X \sim \text{Geom}(p)$*:*
  - $\Pr(X = k) = (1 - p)^{k-1}p$, $k \geq 1$.
  - *Ett experiment upprepas tills händelsen A med sannolikheten p har inträffat och X är antalet upprepningar.*
  - $E(X) = \frac{1}{p}$ *och* $\text{Var}(X) = \frac{1-p}{p^2}$.
- *Negativ binomialfördelning* $X \sim \text{NegBin}(p, r)$, $0 < p \leq 1$, $r \geq 1$*:*
  - *Ett experiment upprepas tills händelsen A med sannolikheten p har inträffat r gånger och X är antalet upprepningar, dvs. X är summan av r oberoende* Geom(p)*-fördelade slumpvariabler.*
  - $\Pr(X = n) = \binom{n-1}{r-1} p^r (1 - p)^{n-r}$.
  - $E(X) = \frac{r}{p}$, *och* $\text{Var}(X) = \frac{r(1-p)}{p^2}$.

Några viktiga kontinuerliga slumpvariabler och deras fördelningar

- *Likformig kontinuerlig fördelning (där $-\infty < a < b < \infty$):*

$$f_X(t) = \begin{cases} \frac{1}{b-a}, & a \le t \le b, \\ 0, & t < a \quad \textit{eller} \quad t > b. \end{cases}$$

$\mathrm{E}(X) = \frac{1}{2}(a+b)$ *och* $\mathrm{Var}(X) = \frac{1}{12}(b-a)^2$.

- *Normalfördelning $X \sim N(\mu, \sigma^2)$:*

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

$\mathrm{E}(X) = \mu$, $\mathrm{Var}(X) = \sigma^2$, *och* $F_{\mathrm{N}(\mu,\sigma^2)}(t) = F_{\mathrm{N}(0,1)}\left(\frac{t-\mu}{\sigma}\right)$.

- *Exponentialfördelning $X \sim \mathrm{Exp}(\lambda)$, $\lambda > 0$:*

$$f_X(t) = \begin{cases} \lambda \mathrm{e}^{-\lambda t}, & t \ge 0, \\ 0, & t < 0, \end{cases} \qquad F_X(t) = 1 - \mathrm{e}^{-\lambda t}, \quad t \ge 0.$$

$E(X) = \dfrac{1}{\lambda}$ *och* $\mathrm{Var}(X) = \dfrac{1}{\lambda^2}$.

Observera att som parameter ofta används väntevärdet $\mu = \frac{1}{\lambda}$.

Sambandet mellan Poisson- och exponentialfördelningen

- *Kunder anländer till en servicepunkt med oberoende och* Exp(λ)*-fördelade intervall om och endast om antalet kunder som kommer inom ett intervall med längden T är en slumpvariabel som har en* Poisson(λT)*-fördelning och antalet kunder som anländer inom disjunkta tidsintervall är oberoende.*
- *I detta fall är väntevärdet längden av tidsintervallet mellan två ankomsttider $\frac{1}{\lambda}$ och väntevärdet av antalet kunder som anländer inom ett tidsintervall med längden T är λT.*
- *Om $\ldots < T_{-1} < T_0 < T_1 < T_2 < \ldots$ så gäller*
  - $U_{(a,b]} = |\{\, j : T_j \in (a,b]\,\}|$ *är* Poisson($\lambda(b-a)$)*-fördelad då $a < b$ och $U_{(a_1,b_1]}$ och $U_{(a_2,b_2]}$ är oberoende om $(a_1,b_1] \cap (a_2,b_2] = \emptyset$*

  om och endast om
  - $T_{j+1} - T_j$ *oberoende och* Exp(λ)*-fördelade för alla j.*

Summan av två oberoende slumpvariabler mm.

- *Antag att X och Y är två oberoende slumpvariabler så att X har täthetsfunktionen f_X och Y har fördelningsfunktionen F_Y. (Motsvarande resultat gäller också då X är diskret.)*
- *Om $a \leq b$ och $A(s) \leq B(s)$ för alla $s \in \mathbb{R}$ så är*

$$\begin{aligned}\Pr(X \in (a,b],\ Y \in (A(X), B(X)]) &= \int_a^b f_X(s) \Pr(Y \in (A(s), B(s)])\,\mathrm{d}s \\ &= \int_a^b f_X(s) \big(F_Y(B(s)) - F_Y(A(s))\big)\,\mathrm{d}s.\end{aligned}$$

- *Slumpvariabelns $X+Y$ fördelningsfunktion är*

$$\begin{aligned}F_{X+Y}(t) &= \Pr(X+Y \leq t) = \Pr(X \in (-\infty,\infty),\ Y \leq t-X) \\ &= \int_{-\infty}^{\infty} f_X(s) \Pr(Y \leq t-s)\,\mathrm{d}s = \int_{-\infty}^{\infty} f_X(s) F_Y(t-s)\,\mathrm{d}s.\end{aligned}$$

- *Om Y har täthetsfunktionen f_Y så har $X+Y$ täthetsfunktionen*

$$f_{X+Y}(t) = \int_{-\infty}^{\infty} f_X(s) f_Y(t-s)\,\mathrm{d}u.$$

Centrala gränsvärdessatsen

Ifall slumpvariablerna $X_1, X_2, \ldots$ *är oberoende och har samma fördelning så att* $\mathrm{E}(X_j) = \mu$ *och* $\mathrm{Var}(X_j) = \sigma^2$, $j = 1, 2, \ldots$, *så gäller*

$$\frac{\frac{1}{n}\sum_{j=1}^{n} X_j - \mu}{\sqrt{\frac{\sigma^2}{n}}} = \frac{\sum_{j=1}^{n} X_j - n\mu}{\sqrt{n\sigma^2}} \sim_a \mathrm{N}(0,1) \quad \text{då } n \to \infty,$$

dvs.

$$\lim_{n\to\infty} \Pr\left(\frac{\sum_{j=1}^{n} X_j - n\mu}{\sqrt{n\sigma^2}} \le t\right) = F_{\mathrm{N}(0,1)}(t) \overset{\text{def}}{=} \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}s^2}\,\mathrm{d}s.$$

Normalapproximation

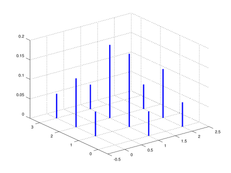
Om X *är summan av "tillräckligt" många oberoende slumpvariabler med ändlig varians så är* $\dfrac{X - \mathrm{E}(X)}{\sqrt{\mathrm{Var}(X)}}$ *ungefär* $\mathrm{N}(0,1)$*-fördelad.*

Tvådimensionella slumpvariabler och fördelningar

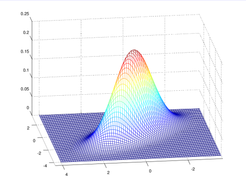
- *Ifall Ω är ett utfallsrum på vilket en sannolikhetsfunktion är definierad så är $(X, Y) : \Omega \to \mathbb{R}^2$ en tvådimensionell slumpvariabel med fördelningsfunktion* $F_{XY}(s, t) = \Pr(X \le s, Y \le t)$ *förutsatt att*

  $\{\, \omega \in \Omega : X(\omega) \le s, Y(\omega) \le t \,\}$ *är en mängd för vilken sannolikheten är definierad.*

- *En tvådimensionell slumpvariabel* (X, Y) *är diskret och den har frekvensfunktionen* $f_{XY}(a, b)$ *ifall* $f_{XY}(a, b) \ge 0$ *för alla a och b,* $\sum_a \sum_b f_{XY}(a, b) = 1$ *och* $\Pr(X = a, Y = b) = f_{XY}(a, b)$ *för alla a och b.*



- *En tvådimensionell slumpvariabel* (X, Y) *är kontinuerlig och har täthetsfunktion* $f_{XY}(s, t)$ *ifall* $f_{XY}(s, t) \ge 0$ *för alla s och t,* $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(s, t)\,\mathrm{d}s\,\mathrm{d}t = 1$ *och* $F_{XY}(s, t) = \int_{-\infty}^{s} \int_{-\infty}^{t} f_{XY}(u, v)\,\mathrm{d}u\,\mathrm{d}v$ *för alla s och t.*



Marginalfördelningar

- *Ifall* $f_{XY}(a,b)$ *är frekvensfunktionen för* *den diskreta tvådimensionella slumpvariabeln* (X,Y) *så är marginalfrekvensfunktionerna för slumpvariablerna* X *och* Y

$$f_X(a) = \Pr(X=a) = \sum_b f_{XY}(a,b) \quad \textit{och}$$

$$f_Y(b) = \Pr(Y=b) = \sum_a f_{XY}(a,b).$$

- *Ifall* $f_{XY}(s,t)$ *är täthetsfunktionen för* *den kontinuerliga tvådimensionella slumpvariabeln* (X,Y) *så är marginaltäthetsfunktionerna för slumpvariablerna* X *och* Y

$$f_X(s) = \int_{-\infty}^{\infty} f_{XY}(s,t)\,\mathrm{d}t \quad \textit{och} \quad f_Y(t) = \int_{-\infty}^{\infty} f_{XY}(s,t)\,\mathrm{d}s.$$

Väntevärden

Om (X, Y) *är en tvådimensionell slumpvariabel och* $h(x, y)$ *är en* *mätbar* *funktion så är*

$$\mathrm{E}\big(h(X, Y)\big) = \sum_a \sum_b h(a, b) f_{XY}(a, b),$$

då (X, Y) *är en diskret slumpvariabel och summan existerar och*

$$\mathrm{E}\big(h(X, Y)\big) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(s, t) f_{XY}(s, t) \, \mathrm{d}s \, \mathrm{d}t,$$

då (X, Y) *är en kontinuerlig slumpvariabel med tähetsfunktion och integral existerar.*

Oberoende slumpvariabler

- *Om* (X, Y) *är en tvådimensionell slumpvariabel (diskret eller kontinuerlig med täthetsfunktion) så gäller*

$$X \textit{ och } Y \textit{ är oberoende} \quad \Leftrightarrow \quad f_{XY}(x,y) = f_X(x) f_Y(y).$$

- *Om* X *och* Y *är oberoende och* f *och* g *är* *mätbara funktioner så gäller*

$$\mathrm{E}(f(X)g(Y)) = \mathrm{E}(f(X))\mathrm{E}(g(Y)).$$

Kovarians och korrelation

- *Kovarians då* $\mathrm{Var}(X) < \infty$ *och* $\mathrm{Var}(Y) < \infty$*:*
  $\mathrm{Cov}(X,Y) = \mathrm{E}\big((X - \mathrm{E}(X))(Y - \mathrm{E}(Y))\big) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y).$
- *Korrelation eller korrelationskoefficient:*
  $$\mathrm{Cor}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}, \ -1 \le \mathrm{Cor}(X,Y) \le 1$$
  då $0 < \mathrm{Var}(X) < \infty$ *och* $0 < \mathrm{Var}(Y) < \infty$.
- *Om* X *och* Y *är oberoende är* $\mathrm{Cor}(X,Y) = 0$ *men om korrelationen är 0 så är* X *och* Y *inte nödvändigtvis oberoende, om inte* (X,Y) *är normalfördelad.*

Betingade fördelningar

- *Ifall* (X, Y) *är en diskret tvådimensionell slumpvariabel med frekvensfunktion* $f_{XY}(a, b)$ *eller en kontinuerlig tvådimensionell slumpvariabel med täthetsfunktion* $f_{XY}(s, t)$ *så är*

$$f_{Y|X}(b|a) = \frac{f_{XY}(a, b)}{f_X(a)}, \quad \textit{då} \quad f_X(a) > 0,$$

  den betingade frekvensfunktionen respektive täthetsfunktionen för Y *givet* $X = a$.

- $\mathrm{E}(Y|X = a) = \begin{cases} \sum_b b f_{Y|X}(b|a) & \textit{eller} \\ \int_{-\infty}^{\infty} t f_{Y|X}(t|a)\,\mathrm{d}t \end{cases}$

- $\mathrm{E}(Y|X)$ *är en slumpvariabel så att* $\mathrm{E}(Y|X)(\omega) = \mathrm{E}(Y|X = X(\omega))$ *då* $\omega \in \Omega$.

- $\mathrm{E}(\mathrm{E}(Y|X)) = \mathrm{E}(Y)$.

Den tvådimensionella normalfördelningen

- *Slumpvariabeln* (X, Y) *är normalfördelad om varje linjärkombination* $\alpha X + \beta Y$ *är normalfördelad.*
- *Om* (X, Y) *är normalfördelad så är också* X *och* Y *normalfördelade, dvs.* $X \sim \mathrm{N}(\mu_X, \sigma_X^2)$ *och* $Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$.
- *Om* (X, Y) *är normalfördelad och* $\mathrm{Cov}(X, Y) = 0$ *(så att också* $\rho_{XY} = \mathrm{Cor}(X, Y) = 0$*) så är* X *och* Y *oberoende.*
- *Om* (X, Y) *är normalfördelad,* $\sigma_X^2 > 0$, $\sigma_Y^2 > 0$ *och* $\rho_{XY} \neq \pm 1$ *så är*

$$f_{Y|X}(t|s) = \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho_{XY}^2}} e^{-\frac{1}{2}\left(\frac{t-\mu_Y-\rho_{XY}\frac{\sigma_Y}{\sigma_X}(s-\mu_X)}{\sigma_Y\sqrt{1-\rho_{XY}^2}}\right)^2} \quad \textit{dvs.}$$

$$(Y|X = s) \sim \mathrm{N}\left(\mu_Y + \rho_{XY}\frac{\sigma_Y}{\sigma_X}(s - \mu_X), (1 - \rho_{XY}^2)\sigma_Y^2\right),$$

och motsvarande resultat gäller för $(X|Y = t)$.

- *Ifall* (X, Y) *är normalfördelad,* $\sigma_X^2 > 0$, $\sigma_Y^2 > 0$ *och* $\rho_{XY} \neq \pm 1$ *så har* (X, Y) *täthetsfunktionen*

$$f_{XY}(t, u) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} e^{-\frac{1}{2(1-\rho_{XY}^2)}\left(\frac{(t-\mu_X)^2}{\sigma_X^2} + \frac{(u-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho_{XY}(t-\mu_X)(u-\mu_Y)}{\sigma_X\sigma_Y}\right)}$$

Obs!

- *Om* (X, Y) *är en slumpvariabel så att* X *är normalfördelad och* Y *är normalfördelad så är inte slumpvariabeln* (X, Y) *nödvändigtvis normalfördelad.*
- *Om* X *och* Y *är oberoende normalfördelade slumpvariabler så är* (X, Y) *normalfördelad.*
- *Om* $X_j \sim \mathrm{N}(\mu_j, \sigma_j^2)$, $j = 1, \ldots, n$ *är oberoende så är* $\sum_{j=1} c_j X_j \sim \mathrm{N}\left(\sum_{j=1}^n c_j \mu_j, \sum_{j=1}^n c_j^2 \sigma_j^2\right)$.

"Tillbaka till väntevärdet"

Om (X, Y) *är normalfördelad så att* X *och* Y *har samma fördelning* $\mathrm{N}(\mu, \sigma^2)$ *och korrelationskoefficienten* $\rho_{XY} = \mathrm{Cor}(X, Y) \neq \pm 1$ *så gäller*

$$\left|\mathrm{E}(Y|X = s) - \mu\right| = |\rho_{XY}||s - \mu| < |s - \mu|.$$