Kombinatorik

$n!$ permutationer/ordningar av en mängd med n element

$\frac{n!}{(n-k)!}$ ordnade urval av k stycken

$\binom{n}{k} = \frac{n!}{k!(n-k)!}$ oordnade urval av k stycken

$\frac{n!}{n_1! n_2! \dots n_k!}$ ordningar av n element där n_j stycken är av typ j och $n = n_1 + \dots + n_k$

Sannolikheter / händelser

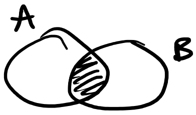
händelse, utfallsrum

$A \subseteq S$

parvis disjunkta

① $\mathbb{P}(A) \geq 0$ ② $\mathbb{P}(S) = 1$ ③ $\mathbb{P}\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mathbb{P}(A_j)$ om $A_i \cap A_j = \emptyset$ för alla $i \neq j$

t.ex. Laplace experiment: $|S| \in \mathbb{N}$, $\mathbb{P}(A) = \frac{|A|}{|S|}$



Räkneregler:

$\mathbb{P}(A') = 1 - \mathbb{P}(A)$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

[$= \mathbb{P}(A) + \mathbb{P}(B)$ om A, B disjunkta]

Betingad slh:

För händelser A, B med $\mathbb{P}(B) > 0$ är $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$

$\Rightarrow \mathbb{P}(A \cap B) = \mathbb{P}(B) \cdot \mathbb{P}(A|B)$

$= \mathbb{P}(A) \cdot \mathbb{P}(B) \Leftrightarrow A, B$ oberoende

Total slh.

(A_j) partition dvs. $\bigcup_j A_j = S$, $A_i \cap A_j = \emptyset \Rightarrow \mathbb{P}(B) = \sum_j \mathbb{P}(B \cap A_j)$

Bayes (A_j) partition, $\mathbb{P}(B) > 0$ med $\mathbb{P}(A_j) > 0$ $\Rightarrow \mathbb{P}(A_j | B) = \frac{\mathbb{P}(B|A_j) \cdot \mathbb{P}(A_j)}{\sum_i \mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)}$

Fördelningar

diskret

Slh.funktion $f_X(x) = \mathbb{P}(X = x)$ för $x \in \mathbb{R}$.

$\mathbb{E}h(X) := \sum_{x \in X(S)} h(x) \cdot \underbrace{f_X(x)}_{\mathbb{P}(X=x)}$

kontinuerlig

täthetsfunktion $f_X(x)$ som uppfyller $\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x)\,dx$ för godtyckliga $a, b \in \mathbb{R}$.

$\mathbb{E}h(X) := \int_{-\infty}^{+\infty} h(x) \cdot f_X(x)\,dx$

Fördelningsfkt. $F_X(x) = \mathbb{P}(X \leq x)$ (för kontinuerlig X: $F_X'(x) = f_X(x)$ där $f_X(x)$ är kontinuerlig)

varians: $\operatorname{var}(X) = \mathbb{E}((X - \mathbb{E}X)^2) = \mathbb{E}X^2 - (\mathbb{E}X)^2$

standardavvikelse: $\sigma = \sqrt{\operatorname{var}(X)}$

Räkneregler: X, Y s.v. $c \in \mathbb{R}$

- $\mathbb{E}(c \cdot X + Y) = c \cdot \mathbb{E}X + \mathbb{E}Y$ (linjäritet)
- $\text{var}(cX) = c^2 \cdot \text{var}(X)$
- $\text{var}(X+Y) = \text{var}(X) + \text{var}(Y)$ bara om X, Y oberoende

moment genererande funktion

derivata av ordning k

$m_X(t) = \mathbb{E}(e^{tX}) \quad \Rightarrow \quad m_X^{(k)}(0) = \mathbb{E}(X^k)$ för godtyckligt $k \in \mathbb{N}$.

Viktiga fördelningar

diskret

- Bernoulli, $\text{Ber}(p)$: $\mathbb{P}(X=1) = p = 1 - \mathbb{P}(X=0)$, $\mathbb{E}X = p$, $\text{var}(X) = p \cdot (1-p)$, $m_X(t) = (1-p+pe^t)$
- binomial, $\text{Bin}(n,p)$: $f_X(k) = \binom{n}{k} p^k (1-p)^{n-k}$ för $k = 0, \dots, n$
  $\mathbb{E}X = np$, $\text{var}(X) = np(1-p)$, $m_X(t) = (1-p+pe^t)^n$.
- geometrisk, $\text{Geom}(p)$: $f_X(k) = p(1-p)^{k-1}$ för $k \in \mathbb{N}$, $F_X(k) = 1-(1-p)^k$
  $\mathbb{E}X = \frac{1}{p}$, $\text{var}(X) = \frac{1-p}{p^2}$, $m_X(t) = \frac{pe^t}{1-(1-p)e^t}$ $[t < -\ln(1-p)]$.
- negativ binomial, $\text{NB}(r,p)$: $f_X(k) = \binom{k-1}{r-1}(1-p)^{k-r} p^r$
  $\mathbb{E}X = \frac{r}{p}$, $\text{var}(X) = \frac{r(1-p)}{p^2}$, $m_X(t) = \left(\frac{pe^t}{1-(1-p)e^t}\right)^r$
- hypergeometrisk, $\text{HypGeo}(N, r; n)$: $f_X(k) = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}$ för $\max\{0, n-(N-r)\} \leq k \leq \min\{n, r\}$
  $\mathbb{E}X = \frac{nr}{N}$, $\text{var}(X) = \frac{nr(N-r)(N-n)}{N^2(N-1)}$
- Poisson, $\text{Poiss}(\lambda)$: $f_X(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$
  $\mathbb{E}X = \lambda$, $\text{var}(X) = \lambda$, $m_X(t) = e^{\lambda(e^t-1)}$

kontinuerlig

- normal, $\mathcal{N}(\mu, \sigma^2)$: $f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$
  $\mathbb{E}X = \mu$, $\text{var}(X) = \sigma^2$, $m_X(t) = \exp\left(\mu t + \frac{(\sigma t)^2}{2}\right)$.
- Gamma, $\Gamma(\alpha, \beta)$: $f_X(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}$ för $x > 0$, där $\Gamma(\alpha) = \int_0^\infty z^{\alpha-1} e^{-z}\, dz$
  $\mathbb{E}X = \alpha\beta$, $\text{var}(X) = \alpha\beta^2$, $m_X(t) = (1-\beta t)^{-\alpha}$ $[t < \frac{1}{\beta}]$.
- exponential, $\text{Exp}(\lambda) = \Gamma(1, \frac{1}{\lambda})$: $f_X(x) = \lambda e^{-\lambda x}$ för $x > 0$, $F_X(x) = 1 - e^{-\lambda x}$
  $\mathbb{E}X = \frac{1}{\lambda}$, $\text{var}(X) = \frac{1}{\lambda^2}$, $m_X(t) = \frac{\lambda}{\lambda - t}$ $[t < \lambda]$.
- chi-kvadrat med γ frihetsgrader, $\chi^2(\gamma) = \Gamma(\frac{\gamma}{2}, 2)$:
  $\mathbb{E}X = \gamma$, $\text{var}(X) = 2\gamma$, $m_X(t) = (1-2t)^{-\gamma/2}$ $[t < \frac{1}{2}]$.
- t-fördelning med γ frihetsgrader: $f_T(t) = \frac{\Gamma(\frac{\gamma+1}{2})}{\Gamma(\frac{\gamma}{2})\sqrt{\gamma\pi}} \left(1 + \frac{t^2}{\gamma}\right)^{-\frac{\gamma+1}{2}}$
  $\mathbb{E}T = 0$ $(\gamma > 1)$, $\text{var}(T) = \frac{\gamma}{\gamma-2}$ $(\gamma > 2)$

Viktiga relationer

- $X \sim Bin(n,p)$ är en summa av n oberoende $Ber(p)$ - variabler.
- $X \sim NB(r,p)$ —— " —— r oberoende $Geom(p)$ - variabler.
- $X \sim Bin(n,p)$ kan approximeras med $Y \sim N(np, np(1-p))$ för fast $p \in (0,1)$ och n stor

  tum regel: $\min\{np, n(1-p)\} > 5$

  Ofta räknas man $\underset{\text{diskret}}{\mathbb{P}(15 \le X < 18)} \approx \underset{\text{kont.}}{\mathbb{P}(14{,}5 \le Y \le 17.5)}$
- $X \sim Bin(n,p)$ är approximativt $Poiss(np)$ fördelat för stort n, litet p (Poisson-antagande)
- $X \sim N(\mu, \sigma^2)$, så är $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$, dvs. standardnormalfördelat
- $X_1, X_2, \dots, X_n \sim N(0,1)$ oberoende $\Rightarrow X_1^2 + X_2^2 + \dots + X_n^2 \sim \chi^2(n)$
- Om $N(t)$ räknar händelser under Poisson-antagande, dvs. $N(t) \sim Poiss(\lambda t)$, så är väntetiden till nästa händelse $Exp(\frac{1}{\lambda})$ - fördelad.
- Om f_X är täthetsfunktionen till X, $Y = h(X)$ för $h: \mathbb{R} \to \mathbb{R}$ strängt monoton och deriverbar, så ges täthetsfunktionen till Y av $f_Y(y) = f_X(h^{-1}(y)) \cdot |(h^{-1})'(y)|$
- $Y = aX + b$ för $a, b \in \mathbb{R} \Rightarrow m_Y(t) = e^{bt} \cdot m_X(at)$
- X, Y oberoende $\Rightarrow m_{X+Y}(t) = m_X(t) \cdot m_Y(t)$

Då momentgenererande funktioner är karakteriserande följer omedelbart:

- Om X_1 och X_2 är oberoende gäller:

  $X_1 \sim Poiss(\lambda_1),\ X_2 \sim Poiss(\lambda_2) \Rightarrow X_1 + X_2 \sim Poiss(\lambda_1 + \lambda_2)$

  $X_1 \sim N(\mu_1, \sigma_1^2),\ X_2 \sim N(\mu_2, \sigma_2^2) \Rightarrow X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

  $X_1 \sim \Gamma(\alpha_1, \beta),\ X_2 \sim \Gamma(\alpha_2, \beta) \Rightarrow X_1 + X_2 \sim \Gamma(\alpha_1 + \alpha_2, \beta)$

  i synnerhet: $X_1 \sim \chi^2(\gamma_1),\ X_2 \sim \chi^2(\gamma_2) \Rightarrow X_1 + X_2 \sim \chi^2(\gamma_1 + \gamma_2)$

  $X_1, X_2 \sim Exp(\lambda) \Rightarrow X_1 + X_2 \sim \Gamma(2, \frac{1}{\lambda})$

- Om $Z \sim N(0,1)$ och $Y \sim \chi^2(\gamma)$ är oberoende $\Rightarrow T = \frac{Z}{\sqrt{Y/\gamma}}$ är t-fördelat med γ frihetsgrader.

Två dimensionella (/ beroende) slumpvariabler (X,Y)

- kan också vara diskreta eller kontinuerliga

$f_{XY}(x,y) = \mathbb{P}(X=x, Y=y)$

täthets flt. $f_{XY}(x,y)$ uppfyller (för $a,b,c,d \in \mathbb{R}$):

$$\mathbb{P}(X\in[a,b], Y\in[c,d]) = \int_a^b \int_c^d f_{XY}(x,y)\,dy\,dx$$

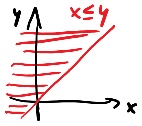
- marginella slh./täthets funktioner

$$f_X(x) = \sum_{y\in Y(S)} f_{XY}(x,y) \qquad f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x,y)\,dy$$

- Gemensamma händelser beräknas genom summation / integration över motsvarande område:

t.ex. $P(X\in A, Y\in B) = \sum_{x\in A}\sum_{y\in B} f_{XY}(x,y)$ $\qquad \mathbb{P}(X\in A, Y\in B) = \int_A \int_B f_{XY}(x,y)\,dy\,dx$

eller $\mathbb{P}(X\le Y) = \sum_{y\in Y(S)} \sum_{x\le y} f_{XY}(x,y)$ $\qquad \mathbb{P}(X\le Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{y} f_{XY}(x,y)\,dx\,dy = \int_{-\infty}^{\infty}\int_{x}^{\infty} f_{XY}(x,y)\,dy\,dx$



- $\mathbb{E}\,h(X,Y) = \sum_{x\in X(S)}\sum_{y\in Y(S)} h(x,y) f_{XY}(x,y)$ $\qquad \mathbb{E}\,h(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y) f_{XY}(x,y)\,dy\,dx$

för $h: \mathbb{R}^2 \to \mathbb{R}$.

- X,Y oberoende $\iff$ $f_{XY}(x,y) = f_X(x)\cdot f_Y(y)$ $\quad$ ($\Rightarrow$ $\mathbb{E}(XY) = \mathbb{E}X\cdot\mathbb{E}Y$ för X,Y oberoende)

- Kovarians: $\operatorname{cov}(X,Y) = \mathbb{E}\left[(X-\mathbb{E}X)(Y-\mathbb{E}Y)\right] = \mathbb{E}(X\cdot Y) - \mathbb{E}X\cdot\mathbb{E}Y$

  cov är symmetrisk och linjär i båda argument.

- Korrelation $\rho_{XY} = \dfrac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)\cdot\operatorname{var}(Y)}}$

  X,Y okorrelerade om $\rho_{XY} = 0$ ($\iff$ $\operatorname{cov}(X,Y)=0$)

  $-1 \le \rho_{XY} \le 1$ och $|\rho_{XY}| = 1 \iff Y = aX+b$ för några $a\ne 0$, $b\in\mathbb{R}$.

- Betingad fördelning:

  slh. funktionen / täthetsfunktionen för Y givet $X=x$ ges av

$$f_{Y|x}(y) = \frac{f_{XY}(x,y)}{f_X(x)}$$