Probability Theory

Inclusion-Exclusion

- $|A \cap B| = |A| + |B| - |A \cap B|$

Permutations: $k!$; order is imp.

Combinations: $\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$

$P(A \cup B) = P(A) + P(B)$ for $\forall$ disjoint A, B

$P^c(A) = 1 - P(A)$ $P(\emptyset) = 0$

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

Conditional Probability

- "probability mass function" = a function giving the prob. of each event

$\sum p_i = 1$ ← $P(A)$

$P(A|B) = \frac{P(A \cap B)}{P(B)}$, $P_B > 0$

dependant:

$P(A \cap B) = P(B|A) \cdot P(A)$, $P_A > 0$

$P(A \cap B^c) = P(A) \cdot P(B^c)$

$P(A^c \cap B^c) = P(A^c) \cdot P(B^c)$

independant:

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

$P(A \cap C) = P(A) \cdot P(C)$

$P(B \cap C) = P(B) \cdot P(C)$

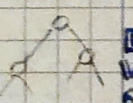
$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$

law of Total Prob.

$P(A) = \sum_{i=1}^{\infty} P(A \cap C_i) = \sum_{i=1}^{\infty} P(A|C_i) \cdot P(C_i)$

Bayes Theorem

$P(R_1|R_2) = \frac{P(R_1 \cap R_2)}{P(R_2)} = \frac{P(R_2|R_1) \cdot P(R_1)}{P(R_2)}$

- 

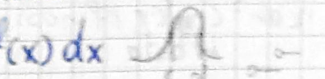
$\mathbb{E}(aX+bY) = \mathbb{E}(X) \cdot a + b \cdot \mathbb{E}(Y)$

$\mathrm{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mu_2$

$\sigma = \sqrt{\mathrm{Var}(X)}$

Random Variables & Distr.

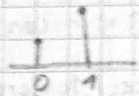
- continous Random Variable:

pdf: $P(c \le x \le d) = \int_c^d f(x)\,dx$ 

cdf: $F(x) = P(X \le x)$ 

$X \sim \mathrm{ber}(p)$

x	0	1
pmf $p(x)$	$1-p$	p
cdf $F(x)$	$1-p$	1



$\mathbb{E}(X) = p$

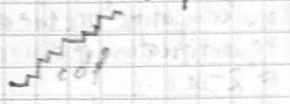
$\mathrm{Var}(X) = p \cdot (1-p)$

$X \sim B(n,p)$ Number of successes in ~~ber(p)~~ n-tries

$B(1,p) = \mathrm{ber}(p)$

$\mathbb{E}(X) = n \cdot p$

$\mathrm{Var} = p \cdot (1-p) \cdot n$



$P(X) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$ ← genau x von n sind richtig

pmf: $\mathrm{dbinom}(x, n, p)$

cdf: $\mathrm{pbinom}(x, n, p, \mathrm{lower.tail} = \mathrm{false}) = \le x$

$X \sim$ geom. Distr. = total Number of attempts before success

$\mathbb{E}(X) = \frac{1-p}{p}$

$\mathrm{Var}(X) = \frac{1-p}{p^2}$

pmf: $p(x) = (1-p)^x \cdot p = \mathrm{dgeom}(n,p)$

cdf: $\mathrm{pgeom}(n,p)$

$X \sim P(\lambda)$ λ = intensity param.

$P(x) = \frac{\lambda^x}{x!} \cdot e^{-\lambda}$, $x \in \{0,1,2,3,\dots\}$: $\mathrm{dpois}(n,\lambda)$

$\mathbb{E}(X) = \mathrm{Var}(X) = \lambda$

$X \sim U(a,b)$

Pdf: $f(x) = \begin{cases} \frac{1}{b-a}, & x \in (a,b) \\ 0, & \text{else} \end{cases}$ $\mathbb{E}(X) = \frac{a+b}{2}$ $\mathrm{Var}(X) = \frac{(b-a)^2}{12}$

cdf: $F(x) = \begin{cases} 0, & x \le a \\ \frac{x-a}{b-a}, & a \le x \le b \\ 1, & x \ge b \end{cases}$

$X \sim \exp(\lambda)$ = models waiting time (continous analouge of gometric distrib.)

pdf: $\begin{cases} \lambda e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$ $\mathbb{E}(X) = \frac{1}{\lambda} = \frac{1}{\lambda}$

cdf: $\begin{cases} 1 - e^{-\lambda x}, & x \ge 0 \\ 0, & x < 0 \end{cases}$ $\mathrm{Var}(X) = \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$

$X \sim N(\mu, \sigma^2)$ Normal distribution

pdf: ~~$\frac{1}{\sqrt{2\pi\sigma^2}}$~~ $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\wedge}\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, $x \in \mathbb{R}$

$\mathbb{E}(X) = \mu$ $X \sim N \to Z \sim N$: $Z = \frac{X-\mu}{\sigma}$

$\mathrm{Var}(X) = \sigma^2$ $Z \sim N \to X \sim N$: $X = \mu + Z\sigma$

pnorm$(\frac{x-\mu}{\sigma})$

Standart Normal distr. $Z \sim N(0,1)$

cdf: $\int_{-\infty}^{z} f(t) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}}\,dt$ $Z = \frac{X-\mu}{\sigma} \Rightarrow P(X > 200) = P\left(\frac{X-\mu}{\sigma} > \frac{200-\mu}{\sigma}\right)$

$P(Z > 2) = \Phi(2.01) = 0.9778\%$

Quantiles:

1σ: ≈ 68%; 2σ: ≈ 95%; 3σ: ≈ 99.7%

0.k Quantile: $z_k = \sigma \cdot \Phi^{-1}(k) + \mu$, $0 < k < 1$

Descriptive Statistics

mean = center of mass of data $\bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i$

(empirical) varience
the mean σ^2 of the data from the mean $s^2 := \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

(empirical) standard deviation $s = \sqrt{s^2}$

Hypothesis Testing

how "compatible" are the data with the assertion μ_0
-> quantification of discrepancy between data & assertion
$\bar{x} - \mu$
$|\;|$ large <-> data hardly compatible with assertion

Z-Tests

Test Statistics: $z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \approx Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$

just measure the discrepancy between data & assertion

p-value: quantifies the discrepancy
$p = \mathbb{P}_{H_0}(|Z| \geq |z|) \approx 9 \cdot 10^{-10}$ example

if p is small, reject Hypothesis
↳ the observed discrepancy was significant

significance level: α

Rejection Area = R

- 2-sided: $H_0: \mu = \mu_0$ ($H_1: \mu \neq \mu_0$)
  $R = (-\infty, q_{\alpha/2}] \cup [q_{1-\frac{\alpha}{2}}, \infty)$
  $P = \mathbb{P}_{H_0}(|Z| \geq |z|)$
- left-sided: $H_0: \mu \geq \mu_0$ ($H_1: \mu < \mu_0$)
  $R = (-\infty, q_{\frac{\alpha}{2}}]$
  $P = \mathbb{P}_{H_0}(Z \leq z)$
- Right-sided: $H_0: \mu \leq \mu_0$ ($H_1: \mu > \mu_0$)
  $R = [q_{1-\frac{\alpha}{2}}, \infty)$
  $P = \mathbb{P}_{H_0}(Z \geq z)$

t-Test (one-sample)

$T := \frac{\bar{x}-\mu}{s/\sqrt{n}} \sim t(n-1)$

n-1 degrees of freedom

$SEM := \frac{s}{\sqrt{n}}$
standart error of mean

mesures the discrepancy $\bar{x}-\mu_0$ in the units SEM

$T = \frac{\bar{x}-\mu_0}{SEM}$

=> if $H_0: \mu = \mu_0$ holds true
$|t| = 1 \Leftrightarrow |\bar{x}-\mu_0| = 1\,sem$
unlikely <- not / very
$|t| = 5 \Leftrightarrow |\bar{x}-\mu_0| = 5\,sem$

Confidence Interval

$I := (\bar{x} - q_{1-\frac{\alpha}{2}} \cdot SEM, \bar{x} + q_{1-\frac{\alpha}{2}} \cdot SEM)$

overlaps the parameter μ_0 with prob. $1-\alpha$

t-Test (two sample)

=> discrepancy lage if SEM small

$t = \frac{\bar{y}-\bar{x}}{\sqrt{sem_y^2 + sem_x^2}}$

$T := \ldots \sim t(r)$ deg. of freedom

$H_0: \mu_x = \mu_y$

$I := ((\bar{y}-\bar{x}) - q_{1-\frac{\alpha}{2}}\sqrt{SEM_y^2 + SEM_x^2},$
$(\bar{y}-\bar{x}) + q_{1-\frac{\alpha}{2}}\sqrt{SEM_y^2 + SEM_x^2})$

Overlaps the parameter μ_0 with prob $1-\alpha$

Proportions

absolute frequency
relative -11- = proportion of $n = \ldots \approx 0$

$H = \frac{1}{n}\sum_{i=1}^{n} Y_i$

$\mathbb{E}(H) = p$ $\quad Var(H) = \frac{p(1-p)}{n}$

$n \to \infty$ $\quad \frac{H-p}{\sqrt{p(1-p)/n}} \to N(0,1)$ (CLT)

$\sigma_H = \sqrt{p \cdot (1-p)/n}$

$SE_H := \sqrt{\frac{H(1-H)}{n}}$ when p is unknown

discrepancy $= |h - p_0| \approx \ldots \cdot se_h$

$Z := \frac{H-p_0}{SE_H} \overset{d}{\approx} N(0,1)$

$I := (H - q_{1-\frac{\alpha}{2}} \cdot se_H, H + q_{1-\frac{\alpha}{2}} \cdot se_H)$

overlaps the parameter p_0 with $1-\alpha$

two sample: $Z := \frac{(H_2 - H_1) - 0}{\sqrt{se^2 + se^2}}$

$I := (H_2 - H_1) - \ldots$
with prob. $1-\alpha$, 0 is overlapped

χ^2-test (goodness of fit)

$X = (X_1 \ldots X_d)^t \sim mult_d(n, p)$

with $p \in (0,1)^d$ & $\sum_{k=1}^{d} p_k = 1$

$H_0: p = (\frac{1}{d} \ldots)^t$

χ^2-stat.: 'observed' -> 'expected'

$\chi^2 := \sum_{k=1}^{d} \frac{(X_k - \mathbb{E}_{H_0}[X_k])^2}{\mathbb{E}_{H_0}[X_k]}$

k	1	2	3	4	5	6	Σ
X_k	21	22	16	17	19	25	120
$\mathbb{E}_{H_0}[X_k]$	20	20	20	20	20	20	120

<- Bsp.