Stochastik für die Naturwissenschaften

Dr. C.J. Luchsinger

Formelsammlung

Fehler od Unvollständigkeit nicht einklagbar bei Prüfung

Enthält eher theoretische Resultate (und nicht komplizierte Teststatistiken, KI's), damit man der Vorlesung besser folgen kann.

Aus 3. Wahrscheinlichkeit P

P: a) $0 \le P[E] \le 1$ für alle E; b) $P[\Omega] = 1$; c) Für endlich oder abzählbar unendlich viele paarweise unvereinbare Ereignisse $E_1, E_2, \ldots$ gilt: $P[E_1 \cup E_2 \cup \ldots] = P[E_1] + P[E_2] + \ldots$.

folgt sofort für unvereinbare E und F: $P[E \cup F] = P[E] + P[F]$.

$A \subseteq B \Rightarrow P[B] = P[A] + P[B \backslash A]$ und damit $A \subseteq B \Rightarrow P[A] \le P[B]$

$P[A \cup B] = P[A] + P[B] - P[A \cap B]$.

$P[A|B] := \frac{P[A \cap B]}{P[B]}$

$P[A|B]P[B] = P[A \cap B] = P[B|A]P[A]$

$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[B|A]P[A]}{P[B|A]P[A] + P[B|\overline{A}]P[\overline{A}]}$

$P[A] = \sum_{i=1}^{n} P[A|B_i]P[B_i]$

A und B (paarweise) unabhängig voneinander, wenn $P[A \cap B] = P[A]P[B]$. Falls $P[B] > 0$ resp. $P[A] > 0$, so ist Unabhängigkeit von A und B gleichbedeutend mit $P[A|B] = P[A]$ resp. $P[B|A] = P[B]$.

Aus 4. Zufallsgrösse X

$F(x) := P[X \leq x] := P[\{\omega | X(\omega) \leq x\}]$, wenn diskret noch $F(x) = P[X \leq x] = \sum_{x_i \leq x} P[X = x_i]$

Y stetig, wenn F_Y sich darstellen lässt als

$$F_Y(a) := P[Y \leq a] = \int_{-\infty}^{a} f(u)du$$

wo $\int_{-\infty}^{\infty} f(x)dx = 1$

stetig: $P[a < X \leq b] = P[X \leq b] - P[X \leq a] = F(b) - F(a) = \int_a^b f(x)dx$

stetig: $P[X = x_0] = \int_{x_0}^{x_0} f(x)dx = 0$

stetig: $P[a \leq X \leq b] = P[a < X \leq b] = P[a < X < b] = P[a \leq X < b]$ $= P[X \leq b] - P[X \leq a] = F(b) - F(a) = \int_a^b f(x)dx$

stetig und F differenzierbar: $F'(x) = f(x)$

Jargon: unabhängig, identisch verteilt = independent and identically distributed (**iid**)

Aus 5. Erwartungswert E und Varianz V

Erwartungswert

$$E[X] := \begin{cases} \sum_{x_i} x_i P[X = x_i] & \text{falls } X \text{ diskret} \\ \int_{-\infty}^{\infty} x f(x)dx & \text{falls } X \text{ stetig} \end{cases}$$

Varianz

$$V[X] := \begin{cases} \sum_{x_i} (x_i - \mu_X)^2 P[X = x_i] & X \text{ diskret} \\ \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x)dx & X \text{ stetig} \end{cases}$$

Standardabweichung sd (**S**tandard **D**eviation) ist die Wurzel aus der Varianz: $sd[X] := \sqrt{V[X]} \geq 0$

Linearität des Erwartungswertes

$E[aX + bY] = aE[X] + bE[Y]$; $E[X + Y] = E[X] + E[Y]$, $E[b] = b$, $E[0] = 0$

$V[aX+b] = a^2 V[X]$ ("Konstante quadratisch raus"), zentral: $V[aX] = a^2 V[X]$

$V[X] = E[X^2] - (E[X])^2$

$$E[X^2] = \begin{cases} \sum_{x_i} x_i^2 P[X = x_i] & \text{falls } X \text{ diskret} \\ \int_{-\infty}^{\infty} x^2 f(x) dx & \text{falls } X \text{ stetig} \end{cases}$$

Varianz einer Summe **bei Unabhängigkeit !**
$V[\sum_{i=1}^n X_i] = \sum_{i=1}^n V[X_i]$ ("Varianz der Summe ist die Summe der Varianzen"); insbesondere $V[X+Y] = V[X] + V[Y]$

Falls X_i iid, sodass $E[X_i] = \mu$ und $V[X_i] = \sigma^2$, dann gelten die folgenden 3 Resulate:

I: $E[\overline{X}] := E\left[\frac{1}{n}\sum_{i=1}^n X_i\right] = \mu$

II: $V[\overline{X}] := V\left[\frac{1}{n}\sum_{i=1}^n X_i\right] = \frac{1}{n}\sigma^2$

III: $sd[\overline{X}] = \frac{1}{\sqrt{n}}\sigma$

Beliebter Blödsinn: $P[X], E[A], E[X = 3]$

Aus 6. Ausgewählte Verteilungen

Diskrete Verteilungen

6.1.1 Bernoulli Be(p)

X kann 2 Werte annehmen: 0 und 1 (alternativ -1 und $+1$). $P[X = 1] = p$ (Erfolg) und $P[X = 0] = 1 - p$ (Misserfolg). $E[X] = p$ und $V[X] = p(1-p)$. Mit $0 \leq k \leq 1$ haben wir

$$P[X = k] = p^k (1-p)^{1-k}$$

6.1.2 Binomial Bin(n,p); Storrer Kapitel 38, 51.1

X_i, $1 \leq i \leq n$, n iid Be(p). $Y := \sum_{i=1}^{n} X_i$. Dann Y per Definitionem Binomialverteilung mit Parametern n und p; Bin(n,p). $E[Y] = np$ und $V[Y] = np(1-p)$. $0 \leq k \leq n$:

$$P[Y = k] = \binom{n}{k} p^k (1-p)^{n-k}$$

Einsatz: Anzahl Erfolge (k) bei n unabhängigen Versuchen mit Erfolgswahrscheinlichkeit p.

6.1.3 Geometrisch Ge(p); Storrer (37.12 b))

X_i, $i \geq 1$, iid Be(p). $Z := \min\{i | X_i = 1\}$. Z Zufallsgrösse auf den natürlichen Zahlen ohne die Null. Z per Definitionem geometrische Verteilung Ge(p) mit Parameter p. Es gilt: $E[Z] = 1/p$ und $V[Z] = (1-p)/p^2$.

$$P[Z = k] = p(1-p)^{k-1}$$

Einzige diskrete Zufallsgrösse, welche *gedächtnislos*: mit $n > m > 0$ gilt

$$P[Z > n | Z > m] = P[Z > (n-m)].$$

Geometrische Verteilung in gewissen Lehrbüchern so definiert, dass man Anzahl Misserfolge bis zum ersten Erfolg zählt. Dann nimmt Ge(p) Werte auf den natürlichen Zahlen *inklusive* 0 an. Resultate analog aber leicht komplizierter.

6.1.4 Poisson Po(λ)

V poissonsch, wenn Sie Werte auf den natürlichen Zahlen inklusive 0 annimmt und zwar mit folgenden Wahrscheinlichkeiten:

$$P[V = v] = e^{-\lambda} \frac{\lambda^v}{v!}.$$

Es gilt: $E[V] = V[V] = \lambda$.

Stetige Verteilungen

6.2.1 Uniform $U[a, b]$

U per Definitionem auf $[a, b]$ uniform verteilt, wenn U folgende Dichtefunktion hat:

$$f(u) = (b - a)^{-1},$$

wobei dann natürlich $a \leq u \leq b$ zu gelten hat. Ausserhalb von $[a, b]$ ist die Dichte gleich null. Es gilt $E[U] = (a + b)/2$ und $V[U] = (b - a)^2/12$.

6.2.2 (Negativ-) Exponential Exp(λ)

Zufallsgrösse X mit Dichtefunktion

$$f(x) = \lambda e^{-\lambda x}, \ \ x \geq 0,$$

heisst exponentialverteilt mit Parameter λ; Exp(λ). $E[X] = 1/\lambda$, $V[X] = 1/\lambda^2$. Modell für: radioaktiver Zerfall, "wann geht Glühbirne kaputt?", Zwischenzeit bei Ankunft von KundInnen in einem Geschäft und vieles mehr.

Einzige stetige Zufallsgrösse, welche *gedächtnislos*:

$$t \geq s, P[X > t|X > s] = P[X > (t - s)].$$

6.2.3 Normal $\mathcal{N}(\mu, \sigma^2)$

Vorsicht: Storrer $\mathcal{N}(\mu, \sigma)$, Luchsinger $\mathcal{N}(\mu, \sigma^2)$

Dichtefunktion

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \ x \in \mathbb{R},$$

$E[X] = \mu$ und $V[X] = \sigma^2$.

"Z-Transform": X sei $\mathcal{N}(\mu, \sigma^2)$, dann $Z := \frac{X-\mu}{\sigma}$ ist $\mathcal{N}(0, 1)$

6.2.4 χ^2

$(X_i)_{i=1}^n$ iid $\mathcal{N}(0,1)$, dann
$$\sum_{i=1}^n X_i^2$$
χ_n^2-verteilt (sprich "Chiquadrat mit n Freiheitsgraden" (degree of freedom, df)).

$(Y_i)_{i=1}^n$ iid $\mathcal{N}(\mu,\sigma^2)$-verteilt, dann
$$\sum_{i=1}^n \frac{(Y_i-\mu)^2}{\sigma^2}$$
χ_n^2-Verteilung. Ein wenig überraschend: Mit
$$S^2 := \sum_{i=1}^n (Y_i - \overline{Y})^2,$$
wo $\overline{Y} := \sum_{i=1}^n Y_i/n$, hat S^2/σ^2 die χ_{n-1}^2-Verteilung. $E[\chi_n^2] = n; V[\chi_n^2] = 2n$.

X sei χ_a^2, Y sei χ_b^2 und X und Y unabhängig voneinander, dann $W := X + Y$ χ_{a+b}^2.

6.2.5 F

U und V zwei unabhängige, χ_m^2-, bzw. χ_n^2-verteilte Zufallsgrössen. Dann
$$W := \frac{U/m}{V/n}$$
F-verteilt mit Parametern m,n: $F_{m,n}$.
$E[W] = n/(n-2)$ (falls $n > 2$).

6.2.6 t

Y eine $\mathcal{N}(0,1)$-Zufallsgrösse und Z eine χ_n^2-Zufallsgrösse; Y unabhängig von Z.
$$T_n := \frac{Y}{\sqrt{Z/n}}$$

Student-t-Verteilung mit n df. Es gilt $E[t_n] = 0$ (falls $n > 1$ und falls $n = 1$ existiert der Erwartungswert nicht), $V[t_n] = n/(n-2)$, sobald $n > 2$. $\sqrt{F_{1,n}} = |t_n|$. X_1, X_2 zwei iid $\mathcal{N}(0,1)$-Zufallgrössen, so ist X_1/X_2 ebenfalls t_1-verteilt. $t_\infty = \mathcal{N}(0,1)$.

Summen von Zufallsgrössen

$$\sum_1^n Be(p) = Bin(n,p)$$

$$\sum_1^n Bin(n_i,p) = Bin(\sum_1^n n_i, p)$$

Summe Normal ist Normal (Summanden müssen nicht mal identisch verteilt sein)

$$\sum_{i=1}^n \mathcal{N}(\mu_i, \sigma_i^2) = \mathcal{N}(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$$

Aus 7. $n \to \infty$ (Konvergenz, LLN, CLT)

CLT: $P\left[\frac{\sum_{k=1}^n X_k - n\mu}{\sqrt{n}\sigma} \leq a\right] \longrightarrow P[\mathcal{N}(0,1) \leq a]$

Aus 8. Schätzen von Parametern und Konfidenzintervalle

$X_1, \ldots, X_n$ iid $\mathcal{N}(\mu, \sigma^2)$, dann hat $\overline{X}$ eine $\mathcal{N}(\mu, \frac{\sigma^2}{n})$-Verteilung

95 % KI f μ bei **bekanntem** σ^2, Daten $x_1, \ldots, x_n$ in Modell $\mathcal{N}(\mu, \sigma^2)$ ist: $[\overline{x} - \frac{1.96\sigma}{\sqrt{n}}, \overline{x} + \frac{1.96\sigma}{\sqrt{n}}]$.

95 % KI f μ bei **unbekanntem** σ^2, Daten $x_1, \ldots, x_n$ in Modell $\mathcal{N}(\mu, \sigma^2)$ ist: $[\overline{x} - \frac{CV\hat{\sigma}}{\sqrt{n}}, \overline{x} + \frac{CV\hat{\sigma}}{\sqrt{n}}]$ wo $\hat{\sigma} := \sqrt{\frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{x})^2}$.

Aus 10. ANOVA/Regression

Auf dem Weg zur Fundamentalgleichung der **Varianzanalyse**:

$\hat{\mu}_j := \bar{Y}_{.j} := \frac{\sum_{i=1}^{n_j} Y_{ij}}{n_j}$

$$GM = \frac{\sum_{j=1}^{k} n_j \bar{Y}_{.j}}{n} = \frac{\sum_{j=1}^{k} n_j \frac{\sum_{i=1}^{n_j} Y_{ij}}{n_j}}{n} = \frac{\sum_{j=1}^{k} \sum_{i=1}^{n_j} Y_{ij}}{n}$$

$\sum_{j=1}^{k} \sum_{i=1}^{n_j} (Y_{ij} - GM)^2 = \sum_{j=1}^{k} \sum_{i=1}^{n_j} (\bar{Y}_{.j} - GM)^2 + \sum_{j=1}^{k} \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2.$

$V := \frac{\sum_{j=1}^{k} n_j (\bar{Y}_{.j} - GM)^2/(k-1)}{\sum_{j=1}^{k} \sum_{i=1}^{n_j} (Y_{ij} - \bar{Y}_{.j})^2/(n-k)}.$

Regression

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i \tag{10.2}$$

$SSE := \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 := \sum_{i=1}^{n} (y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i))^2$

$SSR := \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$

$SS_{xx} := \sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - n\bar{x}^2$

$SS_{yy} := \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - n\bar{y}^2$

$SS_{xy} := \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}$

$SS_{yy} = SSR + SSE$, also

$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}}$ und $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$ sowie $\hat{\sigma}^2 := \frac{1}{n-2} \sum_{i=1}^{n} e_i^2$