Zu Aufgabe 12:

Bayeskriterien in Regrettafeln
Bayes-Entscheidungen hingegen werden durch den Übergang zu Regret nicht beeinflusst.
Bayes-Aktionen:

$$\sum_{j=1}^{m} u(a_i, \vartheta_j)\pi(\{\vartheta_j\}) \longrightarrow \max_{i=1,\dots,n}$$

Übergang zu Regret:

$$\begin{aligned}
& \sum_{j=1}^{m} r(a_i, \vartheta_j)\pi(\{\vartheta_j\}) \longrightarrow \min_{i=1,\dots,n} \qquad \text{(da Verlusttafel)} \\
\iff \quad & \sum_{j=1}^{m} \left(\max_{\ell=1,\dots,n} u(a_\ell, \vartheta_j) - u(a_i, \vartheta_j) \right) \pi(\{\vartheta_j\}) \longrightarrow \min_{i=1,\dots,n} \\
\iff \quad & \sum_{j=1}^{m} \max_{\ell=1,\dots,n} u(a_\ell, \vartheta_j)\pi(\{\vartheta_j\}) - \sum_{j=1}^{m} u(a_i, \vartheta_j)\pi(\{\vartheta_j\}) \longrightarrow \min_{i=1,\dots,n} \\
\iff \quad & (\text{Konst. unabh. von } a_i) - \sum_{j=1}^{m} u(a_i, \vartheta_j)\pi(\{\vartheta_j\}) \longrightarrow \min_{i=1,\dots,n} \\
\iff \quad & \sum_{j=1}^{m} u(a_i, \vartheta_j)\pi(\{\vartheta_j\}) \longrightarrow \max_{i=1,\dots,n}
\end{aligned}$$

d.h. a^* „Bayes-Regret“-Aktion zu $\pi \iff a^*$ Bayes-Aktion zu π.

Zu Aufgabe 13:

$\mathcal{M}$ besteht aus linearen Restriktionen, die sich aus der ordinalen Wahrscheinlichkeitsbewertung ergeben, z.B.

$$\pi(\{\vartheta_{j_1}\}) \geq \pi(\{\vartheta_{j_2}\}) \Longleftrightarrow 0 \leq \pi(\{\vartheta_{j_1}\}) - \pi(\{\vartheta_{j_2}\}),$$

und den trivialen Bedingungen:

$$0 \leq \pi(\{\vartheta_j\}) \leq 1,$$
$$\pi(\{\vartheta_1\}) + \pi(\{\vartheta_2\}) + \pi(\{\vartheta_3\}) = 1,$$

also ist $\mathcal{M}$ in der Tat ein konvexes Polyeder.

Berechnet man die Extremalpunktmenge $\mathcal{E}(\mathcal{M})$, so ergibt sich (siehe z.B. E. Kofler: Prognosen und Stabilität bei unvollständiger Information. Campus (Frankfurt; New York), 1989, p. 26)

$$\mathcal{E}(\mathcal{M}) = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix} \right\}.$$

Damit ist bei a_1

$$\underline{\mathbb{E}}_{\mathcal{M}}(u(a_1)) = \inf_{\pi \in \mathcal{M}} \mathbb{E}_\pi(u(a_1)) = \min_{\pi \in \mathcal{E}(\mathcal{M})} \left(\sum_{j=1}^{m} u(a_1, \vartheta_j)\pi(\{\vartheta_j\}) \right)$$

$$\begin{aligned} &= \min \left(10000 \cdot 1 + 0; 10000 \cdot \frac{1}{2} + 2000 \cdot \frac{1}{2} + 0; \frac{1}{3}(10000 + 20000 - 15000) \right) \\ &= \min (10000; 6000; -1000) = -1000. \end{aligned}$$

und für a_2

$$\underline{\mathbb{E}}_{\mathcal{M}}(u(a_2)) = \inf_{\pi \in \mathcal{M}} \mathbb{E}_\pi(u(a_2)) = \min_{\pi \in \mathcal{E}(\mathcal{M})} \left(\sum_{j=1}^{m} u(a_2, \vartheta_j)\pi(\{\vartheta_j\}) \right)$$

$$\begin{aligned} &= \min \left(1000 \cdot 1; 1000 \cdot \frac{1}{2} + 1000 \cdot \frac{1}{2}; \frac{1}{3}(1000 + 1000 + 0) \right) \\ &= \frac{2000}{3} \end{aligned}$$

$$\underline{\mathbb{E}}_{\mathcal{M}}(u(a_1)) = \min_{\pi \in \mathcal{E}(\mathcal{M})} \mathbb{E}_\pi(u(a_1)) < \min_{\pi \in \mathcal{E}(\mathcal{M})} \mathbb{E}_\pi(u(a_2)) = \underline{\mathbb{E}}_{\mathcal{M}}(u(a_2)).$$

Also ist a_2 Max E Min-Aktion.

Zu Aufgabe 14:

a)

$$X_i \sim N(\mu, \sigma^2) \qquad \vec{X} = (X_1, \ldots, X_n)$$

$$f(x|\mu) = \frac{1}{\sqrt{2\,\pi}^n \cdot \sigma^n} \exp\Big(-\frac{1}{2\,\sigma^2}\sum_{i=1}^n x_i^2 + \frac{2\,\mu}{2\,\sigma^2} n \cdot \bar{x} - \frac{1}{2\,\sigma^2} n \cdot \mu^2\Big)$$

1. $f(x|\mu)$ in Form $f(x|\mu) \propto h(\mu)\ \exp\big(T(x) \cdot b(\mu)\big)$ bringen

$$f(x|\mu) \propto \underbrace{\exp\Big(-\frac{\mu^2 n}{2\sigma^2}\Big)}_{h(\mu)} \cdot \underbrace{\exp\Big(\frac{n\mu\bar{x}}{\sigma^2}\Big)}_{\exp\big(T(x)\cdot b(\mu)\big)}$$

2. Priori der Form $[h(\mu)]^\alpha \exp\big(b(\mu) \cdot \beta\big)$ wählen

$$\pi(\mu) \propto \underbrace{\exp\Big(-\alpha\frac{\mu^2 n}{2\sigma^2}\Big)}_{[h(\mu)]^\alpha} \cdot \underbrace{\exp\Big(\frac{\mu}{\sigma^2}\beta\Big)}_{\exp\big(b(\mu)\beta\big)}$$

Dies ist die zu $X_i \sim N(\mu, \sigma^2), \sigma^2$ bekannt, konjugierte Priori.

Probe: Bestimme Posteriori.

$$\begin{aligned}\pi(\mu|x) \propto h(\mu) \cdot \exp\big(T(x) \cdot b(\mu)\big) \cdot [h(\mu)]^\alpha \cdot \exp\big(b(\mu) \cdot \beta\big) \\ = [h(\mu)]^{\alpha+1} \cdot \exp\big(b(\mu) \cdot (T(x) + \beta)\big)\end{aligned}$$

Dieses Ergebnis entspricht tatsächlich der Form 3.13 und führt genau auf eine Normalverteilung, wie man anhand des folgenden Ansatzes erkennt:

$$\pi(\mu) \propto \exp(-\alpha\frac{\mu^2 n}{2\sigma^2}) \cdot \exp(\frac{\mu}{\sigma^2}\beta) = \exp(-\alpha^* \cdot \mu^2) \cdot \exp(\beta^* \mu)$$

$$\begin{aligned}\pi(\mu) &= \frac{1}{\sqrt{2\,\pi}\rho} \cdot \exp\left(-\frac{1}{2\,\rho^2}(\mu - \nu)^2\right) = \\ &\propto \exp(-\frac{\mu^2}{2\,\rho^2} + \frac{2\,\nu\,\mu}{2\,\rho^2} - \frac{n\nu^2}{2\,\rho^2}) = \\ &\propto \exp(\underbrace{\frac{\nu}{\rho^2}}_{\beta^*} \cdot \mu - \underbrace{\frac{1}{2\,\rho^2}}_{\alpha^*} \cdot \mu^2)\end{aligned}$$

b) $X \sim \text{Binomial}(n,p)$

$$\begin{aligned}
f(x|p) &\propto (1-p)^n \cdot \exp(\ln \frac{p}{1-p}) \\
\pi(p) &\propto (1-p)^{\alpha \cdot n} \cdot \exp(\beta \cdot \ln \frac{p}{1-p}) = \\
&= (1-p)^{\alpha'} \cdot \left(\frac{p}{1-p}\right)^{\beta} = (1-p)^{\alpha'-\beta} \cdot p^{\beta} \quad : \quad \text{Betaverteilung}
\end{aligned}$$

c) $(X_1, \ldots, X_n)$ i.i.d. Poisson-verteilt

$$\begin{aligned}
f(\vec{x}|\vartheta) &= \prod_{i=1}^{n} f(x_i|\vartheta) = \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} \cdot \exp(-\lambda) = \\
&= \frac{\lambda^{\sum_{i=1}^{n} x_i}}{x_1! x_2! \ldots x_n!} \cdot \exp(-n\lambda) \\
&\propto \exp(-n\lambda) \exp(\sum_{i=1}^{n} x_i \ln \lambda) \\
\pi(\lambda) &\propto \exp\big(-(n\lambda) \cdot \alpha\big) \cdot \exp(\beta \ln \lambda) = \\
&= \lambda^{\beta} \cdot \exp(-\alpha' \lambda) \quad : \quad \text{Gammaverteilung}
\end{aligned}$$

Zu Aufgabe 15:

zu a)

$$\widehat{\mu} = \frac{\bar{x}\rho^2 + \nu \cdot \frac{\sigma^2}{n}}{\rho^2 + \frac{\sigma^2}{n}} = \bar{x} \cdot \frac{\rho^2}{\rho^2 + \frac{\sigma^2}{n}} + \nu \cdot \frac{\frac{\sigma^2}{n}}{\rho^2 + \frac{\sigma^2}{n}}$$

zu b) Ceteris paribus Betrachtung

- $\rho^2 \to \infty; \sigma^2, n$ fest

  $$\widehat{\mu} = \bar{x} \cdot \underbrace{\frac{\rho^2}{\rho^2 + \dfrac{\sigma^2}{n}}}_{\to 1} + \nu \underbrace{\frac{\dfrac{\sigma^2}{n}}{\rho^2 + \dfrac{\sigma^2}{n}}}_{\to 0} \to \bar{x} \text{ ggf. durch } \rho^2 \text{ teilen}$$

  $$\uparrow$$

  $$\rho^{2\prime} = \frac{\rho^2 \cdot \dfrac{\sigma^2}{n}}{\rho^2 + \dfrac{\sigma^2}{n}} = \frac{\dfrac{\sigma^2}{n}}{1 + \dfrac{\sigma^2}{\rho^2 \cdot n}} \qquad \rho^2 \to \infty : \frac{\sigma^2}{n} \text{ wie ML}$$

  ρ^2 große Streuung Priori (heißt geringes Vorwissen)

  Stichprobe bekommt großes Gewicht

- $\sigma^2 \to \infty; \rho^2, n$ fest

  $$\widehat{\mu} = \bar{x} \cdot \underbrace{\frac{\rho^2}{\rho^2 + \dfrac{\sigma^2}{n}}}_{\to 0} + \nu \underbrace{\frac{\dfrac{\sigma^2}{n}}{\rho^2 + \dfrac{\sigma^2}{n}}}_{\to 1} \to \nu$$

  σ^2 sehr groß
  Stichprobe sehr stark streuend
  $\Rightarrow$ man bleibt bei Priori-Wissen,
  da man schlecht dazulernen kann

- $n \to \infty; \sigma^2, \rho^2$ fest

$$\widehat{\mu} = \bar{x} \cdot \underbrace{\frac{\rho^2}{\rho^2 + \frac{\sigma^2}{n}}}_{\to 1} + \nu \underbrace{\frac{\frac{\sigma^2}{n}}{\rho^2 + \frac{\sigma^2}{n}}}_{\to 0} \to \bar{x}$$

bei sehr großen Stichproben überwiegen

die Beobachtungen völlig;

Priori Ausgangspunkt spielt keine Rolle mehr

Dieser Aspekt wird später verallgemeinert:

(„asymptotische“ Objektivität)

zu c) Weiterführende Bemerkung:

Konvexkombination zwischen Stichprobenmittel und Priori-Mittelwert bringt auch ein „Verwaschen“ potentieller widersprüchlicher Information!
Sei

$$\varrho^2 = \frac{\sigma^2}{n}$$

so ist

$$\hat{\mu} = \nu' = \frac{\bar{x} + \nu}{2}$$

und

$$\varrho^{2'} = \frac{\varrho^4}{2\varrho^2} = \frac{\varrho^2}{2}$$

.
Beispielsweise

$$\bar{x} = 0.9 \text{ und } \nu = 1.1$$

führen genau auf dieselbe Posteriori wie

$$\bar{x} = -100 \text{ und } \nu = 102.$$

(massiver Priori-Datenkonflikt!) Eigentlich würde man im zweiten Fall viel „unsicherer“ sein wollen als im ersten. Starkes Argument für Intervallwahrscheinlichkeit → aktuelle Forschung (Gero Walter).

Zu Aufgabe 16:

Hat man bereits das Auswertungsproblem bestimmt, so ergibt sich die Laplace-optimal Aktion über die zeilenweise gebildeten Mittelwerte.

Notation: $d(1,2,2) = \begin{pmatrix} x_1 & x_2 & x_3 \\ a_1 & a_2 & a_2 \end{pmatrix}$

Datengestütztes Entscheidungsproblem:

	ϑ_1	ϑ_2	ϑ_3	Minimum	Laplace
$d(1,1,1)$	10 000	2 000	– 15 000	– 15 000	– 1 000
$d(1,1,2)$	9 100	1 600	– 7 500	– 7 500	1 066 $.\overline{6}$
$d(1,2,1)$	7 300	1 600	– 9 000	– 9 000	– 33 $.\overline{3}$
$d(1,2,2)$	6 400	1 200	– 1 500	– 1 500	**2 033** $\mathbf{.\overline{3}}$
$d(2,1,1)$	4 600	1 800	– 13 500	– 13 500	– 2 366 $.\overline{6}$
$d(2,1,2)$	3 700	1 400	– 6 000	– 6 000	– 300
$d(2,2,1)$	1 900	1 400	– 7 500	– 7 500	– 1 400
$d(2,2,2)$	1 000	1 000	0	**0**	666 $.\overline{6}$

Verwendet man den Hauptsatz, so ist für x_s, $s = 1,\ldots,3$, die Posteriori-Verteilung $\pi(\cdot|x_s)$ über $(\Theta, \mathcal{P}(\Theta))$ zu berechnen und dann das ursprüngliche datenfreie Entscheidungsproblem mit $\pi(\cdot|x_s)$ als aufdatierter Priori zu lösen.

$$\pi(\vartheta_j|x_s) = \frac{p(x_s|\vartheta_j)\cdot\pi(\{\vartheta_j\})}{p(x_s|\vartheta_1)\cdot\pi(\{\vartheta_1\}) + p(x_s|\vartheta_2)\cdot\pi(\{\vartheta_2\}) + p(x_s|\vartheta_3)\cdot\pi(\{\vartheta_3\})}$$

z.B.

$$\pi(\vartheta_1|x_1) = \frac{0.6\cdot 1/3}{0.6\cdot 1/3 + 0.2\cdot 1/3 + 0.1\cdot 1/3} = \frac{0.6}{0.9} = \frac{2}{3}$$

Wenn x_1 (positive Konjunkturprognose) beobachtet wird, dann steigt die Wahrscheinlichkeit, dass tatsächlich Konjunktur steigt, von 1/3 auf 2/3.

Damit

$$\begin{pmatrix} \pi(\vartheta_1|x_1) \\ \pi(\vartheta_2|x_1) \\ \pi(\vartheta_3|x_1) \end{pmatrix} = \begin{pmatrix} 2/3 \\ 2/9 \\ 1/9 \end{pmatrix}$$

$$\begin{pmatrix} \pi(\vartheta_1|x_2) \\ \pi(\vartheta_2|x_2) \\ \pi(\vartheta_3|x_2) \end{pmatrix} = \begin{pmatrix} 3/11 \\ 4/11 \\ 4/11 \end{pmatrix}$$

$$\begin{pmatrix} \pi(\vartheta_1|x_3) \\ \pi(\vartheta_2|x_3) \\ \pi(\vartheta_3|x_3) \end{pmatrix} = \begin{pmatrix} 0.1 \\ 0.4 \\ 0.5 \end{pmatrix}$$

Kontrolle: Posteriori-*Wahrscheinlichkeit*, Komponenten müssen sich zu 1 addieren.

	ϑ_1	ϑ_2	ϑ_3	$\begin{pmatrix} 2/3 \\ 2/9 \\ 1/9 \end{pmatrix}$	$\begin{pmatrix} 3/11 \\ 4/11 \\ 4/11 \end{pmatrix}$	$\begin{pmatrix} 0.1 \\ 0.4 \\ 0.5 \end{pmatrix}$
a_1	10000	2000	-15000	$\underline{5444}$	-2000	-5700
a_2	1000	1000	0	889	$\underline{636}$	$\underline{500}$

(gerundete Werte)
damit ist

bei $\begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix}$ die Aktion $\begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix}$ Posteriori-Nutzen-optimal.

Als optimale Entscheidungsfunktion ergibt sich in der Tat

$$\begin{pmatrix} x_1 & x_2 & x_3 \\ a_1 & a_2 & a_2 \end{pmatrix} = d(1, 2, 2),$$

wie aus der Analyse des Priori-Gain.
Hier nur zur Illustration alle x_s „durchdekliniert“. In der Praxis reicht es, das konkret beobachtete x zu betrachten.