1 1

1.1

Če zapišemo gostote v eksponentni obliki, dobimo

$$f(x;a,b) = \frac{b^a}{\Gamma(a)} e^{-\langle (a+1,b),(\ln(x),1/x)\rangle} \cdot \mathbb{1}_{(0,\infty)}(x),$$

torej dobimo po Neuman-Fischerjevem faktorizacijskem izreku, da je

$$T(x) = (\ln(x), 1/x)$$

zadostna statistika za ta model.

1.2

Ker smo v parametričnem modelu z zvezno odvedljivimi gostotami, lahko iščemo stacionarne točke gostot z logaritmično enačbo verjetja. Računamo

$$\begin{aligned}
&l(a,b) := \ln(f(x;a,b)) = a\ln(b) - \ln(\Gamma(a)) - (a+1)\ln(x) - \frac{b}{x}\\
&\frac{\partial l}{\partial a}(a,b) = \ln(b) - \frac{\Gamma'(a)}{\Gamma(a)} - \ln(x)\\
&\frac{\partial l}{\partial b}(a,b) = \frac{a}{b} - \frac{1}{x}.
\end{aligned}$$

Zapišemo lahko logaritemsko enačbo verjetja

$$\ln(b) - \frac{\Gamma'(a)}{\Gamma(a)} - \ln(x) = \frac{a}{b} - \frac{1}{x} = 0,$$

iz česar takoj dobimo $b = ax$, za drugo enačbo pa potem sledi, da rešujemo

$$\Psi(a) := \ln(a) - \frac{\Gamma'(a)}{\Gamma(a)} = 0.$$

Po posledici Bernsteinovega izreka o monotonih funkcijah velja, da je za $x > 0$

$$\Psi(x) - \frac{1}{2x} > 0,$$

torej naša funkcija nima ničel in zato cenilka po metodi največjega verjetja ne obstaja (digamma wiki).

1.3

Za $k \in \mathbb{N}$ velja

$$m_k := P_{(a,b)}[\mathrm{id}^k] = \int_0^\infty \frac{b^a}{\Gamma(a)} x^{-a-1} e^{-b/x} x^k dx = \frac{b^a}{\Gamma(a)} \frac{\Gamma(a-k)}{b^{a-k}} \int_0^\infty \frac{b^{a-k}}{\Gamma(a-k)} x^{-(a-k)-1} e^{-b/x} dx = b^k \frac{\Gamma(a-k)}{\Gamma(a)}.$$

Označimo

$$\begin{aligned}
&\hat{b}(x_1,x_2,x_3) := \left(\frac{x_1}{x_2} - \frac{x_2}{x_3}\right)^{-1}\\
&\hat{a}(x_1,x_2,x_3) := \frac{\hat{b}(x_1,x_2,x_3)}{x_1} + 1
\end{aligned}$$

in opazimo, da je $\hat{b}(m_1,m_2,m_3) = b$ in $\hat{a}(m_1,m_2,m_3) = a$, torej dobimo zaporedje cenilk po metodi momentov

$$(\hat{a},\hat{b}) \circ (\hat{m_1^n}, \hat{m_2^n}, \hat{m_3^n}),$$

kjer je

$$\hat{m_k^n}(x) = \frac{1}{n}\sum_{i=1}^n x_i^k.$$

Ker je $g = (\hat{a},\hat{b})$ zvezna na $\mathbb{R}^3_+$, je ta cenilka dosledna. Sicer pa g ni trivialna preslikava in zato ne moremo enostavno analizirati porazdelitve, niti nujno izračunati povprečne vrednosti.

1.4

Naj bodo torej $X_i \sim P_\vartheta$ NEP. Definirajmo NEP vektorje $Y_i := (X_i, X_i^2, X_i^3)$ in označimo

$$M_n := (\hat{m_1^n}, \hat{m_2^n}, \hat{m_3^n}) \quad \text{in} \quad \hat{\vartheta}_n := g \circ M_n.$$

Po centralnem limitnem izreku velja

$$\sqrt{n} \cdot (M_n(X_1, \ldots, X_n) - \mu) = \sqrt{n} \cdot \left(\frac{1}{n} \sum_{i=1}^{n} Y_i - \mu \right) \xrightarrow[n\to\infty]{D} N(0, \Sigma),$$

kjer je μ pričakovana vrednost, Σ pa variančno-kovariančna matrika vektorjev Y_i. Potem nam metoda delta skupaj s Cramer-Woldovim pravilom zagotovi

$$\sqrt{n} \cdot (\hat{\vartheta}_n(X_1, \ldots, X_n) - \vartheta) = \sqrt{n} \cdot (g(M_n(X_1, \ldots, X_n)) - g(\mu)) \xrightarrow[n\to\infty]{D} N(0, \mathscr{J}_g(\mu)\Sigma\mathscr{J}_g(\mu)^T),$$

saj je očitno $\mu = [m_1, m_2, m_3]^T$.

1.5

Enostavno izračunamo

$$\Sigma = \begin{bmatrix} m_2 - m_1^2 & m_3 - m_1 m_2 & m_4 - m_1 m_3 \\ m_3 - m_1 m_2 & m_4 - m_2^2 & m_5 - m_2 m_3 \\ m_4 - m_1 m_3 & m_5 - m_2 m_3 & m_6 - m_3^2 \end{bmatrix}$$

in označimo

$$A := \mathscr{J}_g(\mu)\Sigma\mathscr{J}_g(\mu)^T.$$

Definirajmo sedaj zaporedje cenilk $\hat{A}_n$ tako, da zamenjamo v A elemente m_i z $\hat{m_i^n}$. Potem uporabimo Sluckyjev izrek in zveznost korena ter inverza spd matrike, da dobimo

$$\pi_n(X_1, \ldots, X_n, \vartheta) := \sqrt{n} \cdot \hat{A}_n^{-1/2} \left(\hat{\vartheta}_n(X_1, \ldots, X_n) - \vartheta \right) \xrightarrow[n\to\infty]{D} N(0, I).$$

Naj bo $B \subset \mathbb{R}^2$ neka množica, ki zadošča

$$P(\pi_n(X_1, \ldots, X_n, \vartheta) \in B) \approx N(0, I)(B) = 1 - \alpha,$$

Sledi

$$P(\pi_n(X_1, \ldots, X_n, \vartheta) \in B) = P(\hat{\vartheta}_n(X_1, \ldots, X_n) - \vartheta \in \hat{A}_n^{1/2}(B)/\sqrt{n}) = P(\vartheta \in \hat{\vartheta}_n(X_1, \ldots, X_n) - \hat{A}_n^{1/2}(B)/\sqrt{n}),$$

torej je območje zaupanja za ϑ enako

$$\hat{\vartheta}_n(X_1, \ldots, X_n) - \hat{A}_n^{1/2}(B)/\sqrt{n}.$$

Vprašanje izbire množice B lahko rešimo s predpostavko, da je to zaprta krogla s središčem v 0 v neki normirani metriki, s čimer se verjetno da dokazati enoličnost njenega radija.

2 2

Naj bodo $X_i \sim P_\mu$ NEP. Po Rao-Blackwellevem izreku lahko najdemo NCEND kot pogojno upanje nepristranske cenilke $\mathbb{1}_{\{X_1 \leqslant 5\}}$ pogojno na kompletno in zadostno statistiko

$$\bar{X}^n = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Označimo $Y = \bar{X}^n - \mu$ in definirajmo

$$\tilde{X} = X_1 - \rho Y, \text{ kjer je } \rho = P[X_1 Y]/\sigma(Y)^2.$$

Zanjo velja

$$P[\tilde{X}Y] = P[X_1 Y] - \rho P[Y^2] = 0,$$

torej sta $\tilde{X}$ in Y nekorelirani, ker pa sta normalni, sta tudi neodvisni. Sedaj lahko zapišemo

$$\begin{aligned}
P[\mathbb{1}_{\{X_1 \leqslant 5\}} \mid \bar{X}^n = t] &= P(X_1 \leqslant 5 \mid \bar{X}^n = t) = P(\tilde{X} + \rho Y \leqslant 5 \mid Y + \mu = t) = P(\tilde{X} + \rho(t-\mu) \leqslant 5) \\
&= P(X_1 - \rho(\bar{X}^n - \mu) + \rho(t-\mu) \leqslant 5) = P(X_1 - \rho \bar{X}^n \leqslant 5 + \rho t).
\end{aligned}$$

Konkretno:

$$P[X_1 Y] = P[(\sigma Z_1 + \mu)Y] = P[\sigma Z_1 Y] = P[\sigma Z_1 \sum_{i=1}^n \frac{\sigma}{n} Z_i] = \frac{\sigma^2}{n} P[Z_1^2] = \frac{\sigma^2}{n},$$

torej skupaj z dejstvom $\sigma(Y) = \frac{\sigma}{\sqrt{n}}$ dobimo $\rho = 1$ in zaporedje cenilk

$$U_n(\bar{x}^n) = P(\sigma(1 + 1/\sqrt{n})Z \leqslant 5 + \bar{x}^n) = \Phi\left(\frac{5 + \bar{x}^n}{\sigma(1 + 1/\sqrt{n})}\right)$$

3 4

3.1 a)

Upoštevamo Neuman-Pearsonova pogoja

$$P_{p_0}[\phi_{p_0}] = \alpha \quad \text{in} \quad \frac{d}{dp_0} P_{p_0}[\phi_{p_0}] = 0.$$

Označimo $T(x_1, \ldots, x_n) = \sum_{i=1}^n x_i$ in izrazimo enačbi:

$$\begin{aligned}
P_{p_0}[\phi_{p_0}] &= P_{p_0}(T < C_1(p_0)) + P_{p_0}(T > C_2(p_0)) + \gamma_1(p_0) P_{p_0}(T = C_1(p_0)) + \gamma_2(p_0) P_{p_0}(T = C_2(p_0)) \\
&= \mathrm{Bin}(n, p_0)(\{0, \ldots, C_1(p_0) - 1, C_2(p_0) + 1, \ldots, n\}) \\
&\quad + \gamma_1(p_0) \mathrm{Bin}(n, p_0)(\{C_1(p_0)\}) + \gamma_2(p_0) \mathrm{Bin}(n, p_0)(\{(C_2(p_0)\}) \\
&= 1 - \sum_{i=C_1(p_0)}^{C_2(p_0)} \binom{n}{i} p_0^i (1-p_0)^{n-i} + \sum_{i \in \{1,2\}} \gamma_i(p_0) \binom{n}{C_i(p_0)} p_0^{C_i(p_0)} (1-p_0)^{n - C_i(p_0)} \\
&= \alpha
\end{aligned}$$

in

$$\begin{aligned}
\frac{d}{dp_0} P_{p_0}[\phi_{p_0}] &= - \sum_{i=C_1(p_0)}^{C_2(p_0)} \binom{n}{i} \left(i p_0^{i-1} (1-p_0)^{n-i} - (n-i) p_0^i (1-p_0)^{n-i-1} \right) \\
&\quad + \sum_{i \in \{1,2\}} \gamma_i(p_0) \binom{n}{C_i(p_0)} \left(C_i(p_0) p_0^{C_i(p_0 - 1)} (1-p_0)^{n - C_i(p_0)} - (n - C_i(p_0)) p_0^{C_i(p_0-1)} (1-p_0)^{n - C_i(p_0) - 1} \right) \\
&= 0.
\end{aligned}$$

Če privzamemo, da so vse vrednosti razen $\gamma_i(p_0)$ konstantne, je to sistem dveh linearnih enačb za $\gamma_i(p_0)$. To, da so ostale stvari konstantne dosežemo tako, da testiramo vsak $C_1(p_0) \in \{0, \ldots, n-1\}$ in $C_2(p_0) \in \{C_1(p_0) + 1, \ldots, n\}$ posebej. Vzamemo prve vrednosti, pri katerih velja $0 \leqslant \gamma_i(p_0) \leqslant 1$ za $i = 1, 2$.

Za interval zaupanja definiramo

$$I(x) := \{p \in (0,1) \mid \phi_p(x) = 0\} = \{p \in (0,1) \mid T(x) \in [C_1(p), C_2(p)]\}.$$

Po definiciji / izpeljavi randomiziranega testa ϕ_p, velja

$$P_p(p \in I) = P_p(\phi_p = 0) \geqslant 1 - P_p[\phi_p] = 1 - \alpha.$$