Solutions to Problem Sheet 4: Bayesian Inference, Tests, and Decision Theory

1. **HPD Regions**

   (a) If a posterior distribution on $\mathbb{R}$ has a continuous, symmetric density that is strictly increasing on $\mathbb{R}^-$ and strictly decreasing on $\mathbb{R}^+$, show that the HPD regions of level $1-\alpha$ coincide with the intervals defined by the quantiles $\alpha/2$ and $1-\alpha/2$ of the posterior density.

   Solution: The posterior density, let us denote it g, being continuous and unimodal, by the intermediate value theorem the equation $g(t) = y$ always has two solutions for all $y \in J$ with $J = (0, \|g\|_\infty)$. By symmetry of g, the level set $\mathcal{H}(y) = \{t : g(t) \geq y\}$ is written as

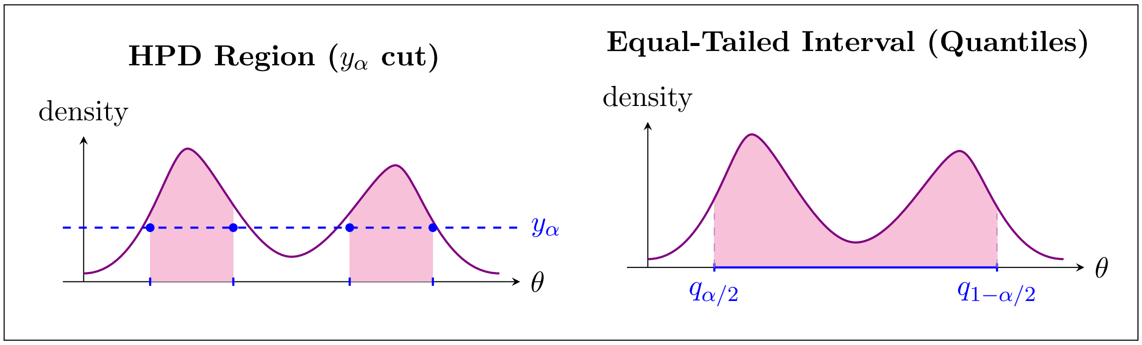
   $$\mathcal{H}(y) = [m - u_y, m + u_y],$$

   where m is such that $g(m) = \|g\|_\infty$, the point where the maximum of the density is reached. We deduce that the HPD region of level $1-\alpha$ is of this form, namely $[m - u_{y_\alpha}, m + u_{y_\alpha}]$, for a well-chosen y_α. Now, the region constructed with the quantiles is also written in this form, by symmetry of g. Since the credibility of each of these regions is the same, namely exactly $1-\alpha$ (since g is continuous), we deduce that they coincide.

   (b) Give an example of a posterior density for which the two types of $1-\alpha$ credible regions from the previous question do not coincide.

   Solution: Any asymmetric (skewed) distribution or multi-modal distribution will serve as a counter-example. Consider a bimodal mixture posterior (e.g., a mixture of two Beta or Gaussian distributions).
   The figure below illustrates the difference:

   - **Left (HPD):** The Highest Posterior Density region is defined by a density threshold y_α (blue dashed line). It selects only the "high probability" peaks, resulting in two *disjoint* intervals. This is the smallest volume set containing probability $1-\alpha$.
   - **Right (Quantile):** The Quantile Credible Interval is defined by cutting off $\alpha/2$ mass from each tail. It results in a single *connected* interval that necessarily includes the low-probability "valley" between the modes.



2. **Credible Intervals and Confidence Intervals**

Let $X_i \sim \text{Bernoulli}(\theta)$ and $\theta \sim \text{Beta}(a, b)$.

(a) Determine the posterior distribution $\Pi[\cdot|X]$. We will denote its mean by m_X and its variance by v_X.

Solution: The Beta distribution is conjugate to the Bernoulli likelihood. The likelihood is $L(\theta) \propto \theta^{\sum X_i}(1-\theta)^{n-\sum X_i}$. The prior is $\pi(\theta) \propto \theta^{a-1}(1-\theta)^{b-1}$. The posterior is proportional to:

$$\pi(\theta|X) \propto \theta^{a+\sum X_i - 1}(1-\theta)^{b+n-\sum X_i - 1}.$$

Thus, the posterior is $\text{Beta}(A, B)$ with $A = a + S_n$ and $B = b + n - S_n$, where $S_n = \sum_{i=1}^n X_i$.

The mean and variance are:

$$m_X = \frac{A}{A+B} = \frac{a+S_n}{a+b+n},$$

$$v_X = \frac{AB}{(A+B)^2(A+B+1)} = \frac{(a+S_n)(b+n-S_n)}{(a+b+n)^2(a+b+n+1)}.$$

(b) Construct a credible interval $I^T(X)$ of level at least $1-\alpha$ (with $\alpha > 0$), centered at m_X, using Chebyshev's inequality.

Solution: Chebyshev's inequality states that for a random variable θ with mean m_X and variance v_X:

$$P(|\theta - m_X| \geq k\sqrt{v_X}) \leq \frac{1}{k^2}.$$

Taking the complement:r

$$P(|\theta - m_X| < k\sqrt{v_X}) \geq 1 - \frac{1}{k^2}.$$

To achieve a credible level of at least $1-\alpha$, we set $1 - 1/k^2 = 1 - \alpha \implies k = 1/\sqrt{\alpha}$. The resulting interval is:

$$I^T(X) = \left[m_X - \frac{1}{\sqrt{\alpha}}\sqrt{v_X}, \quad m_X + \frac{1}{\sqrt{\alpha}}\sqrt{v_X} \right].$$

(c) We ask whether $I^T(X)$ can be used as an asymptotic confidence interval, in the frequentist sense under P_{θ_0}. Answer this question by seeking an asymptotic lower bound for the level of $I^T(X)$ as a function of α.

Solution: Yes, $I^T(X)$ is a valid (and conservative) asymptotic confidence interval. We construct the proof in three steps.

1. Asymptotic Normality of the Posterior Mean

Recall that $m_X = \frac{S_n+a}{n+a+b}$, where $S_n = \sum_{i=1}^n X_i$. We center and scale this estimator around the true parameter θ_0:

$$\sqrt{n}(m_X - \theta_0) = \sqrt{n}\left(\frac{S_n + a}{n+a+b} - \theta_0\right).$$

Rearranging terms:

$$\sqrt{n}(m_X - \theta_0) = \sqrt{n}\left(\frac{S_n - n\theta_0 + a - (a+b)\theta_0}{n+a+b}\right) = \frac{n}{n+a+b}\left(\frac{S_n - n\theta_0}{\sqrt{n}}\right) + \frac{\sqrt{n}(a-(a+b)\theta_0)}{n+a+b}.$$

We analyze the limits as $n \to \infty$:

- By the Central Limit Theorem (CLT), $\frac{S_n - n\theta_0}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, \theta_0(1-\theta_0))$.
- The coefficient $\frac{n}{n+a+b} \to 1$.
- The remainder term $\frac{\sqrt{n}C}{n+a+b} \approx \frac{C}{\sqrt{n}} \to 0$.

By **Slutsky's Theorem**, the sum converges to the distribution of the leading term:

$$\sqrt{n}(m_X - \theta_0) \xrightarrow{d} \mathcal{N}(0, \theta_0(1-\theta_0)). \tag{1}$$

2. Consistency of the Scaled Posterior Variance

Recall $v_X = \frac{(a+S_n)(b+n-S_n)}{(a+b+n)^2(a+b+n+1)}$. We study the behavior of nv_X:

$$nv_X = \frac{n}{(a+b+n+1)} \cdot \frac{a+S_n}{a+b+n} \cdot \frac{b+n-S_n}{a+b+n}.$$

- By the Strong Law of Large Numbers (SLLN), $\frac{S_n}{n} \xrightarrow{a.s.} \theta_0$.
- Thus, $\frac{a+S_n}{a+b+n} = \frac{a/n + S_n/n}{a/n + b/n + 1} \xrightarrow{a.s.} \theta_0$.
- Similarly, $\frac{b+n-S_n}{a+b+n} \xrightarrow{a.s.} 1 - \theta_0$.
- The leading term $\frac{n}{a+b+n+1} \to 1$.

By the Continuous Mapping Theorem (CMT), the product converges almost surely:

$$nv_X \xrightarrow{a.s.} \theta_0(1-\theta_0). \tag{2}$$

3. Convergence of the Pivotal Quantity Z_n

Let Z_n be the standardized deviation used in the credible interval:

$$Z_n = \frac{m_X - \theta_0}{\sqrt{v_X}} = \frac{\sqrt{n}(m_X - \theta_0)}{\sqrt{n v_X}}.$$

From (1), the numerator converges in distribution to $\mathcal{N}(0, \theta_0(1-\theta_0))$. From (2), the denominator converges almost surely (and thus in probability) to $\sqrt{\theta_0(1-\theta_0)}$.
Applying **Slutsky's Theorem** (ratio rule):

$$Z_n \xrightarrow{d} \frac{\mathcal{N}(0, \theta_0(1-\theta_0))}{\sqrt{\theta_0(1-\theta_0)}} \sim \mathcal{N}(0,1).$$

Conclusion
The credible interval condition is $|\theta_0 - m_X| \leq \frac{1}{\sqrt{\alpha}}\sqrt{v_X}$, which is equivalent to $|Z_n| \leq \frac{1}{\sqrt{\alpha}}$. The asymptotic coverage probability is:

$$\lim_{n\to\infty} P_{\theta_0}(\theta_0 \in I^T(X)) = P\left(|\mathcal{N}(0,1)| \leq \frac{1}{\sqrt{\alpha}}\right).$$

By Chebyshev's inequality (used to construct the interval), we know that for any random variable Y, $P(|Y-\mu| < k\sigma) \geq 1 - 1/k^2$. Here $k = 1/\sqrt{\alpha}$, so the lower bound is $1-\alpha$. However, for the Normal distribution, the mass concentration is much higher than the Chebyshev bound implies. For instance, if $\alpha = 0.05$, $1/\sqrt{\alpha} \approx 4.47$, whereas the standard normal requires only 1.96 for 95% coverage.
Thus, the asymptotic level is strictly greater than $1-\alpha$.

3. **Bayesian Testing**

(a) **Test I:** Let $X = (X_1, \ldots, X_n)|\theta \sim \mathcal{N}(\theta, \sigma^2)^{\otimes n}$ and $\theta \sim \Pi = \mathcal{N}(\mu, \tau^2)$, where σ^2, τ^2 are fixed.

i. Determine the posterior distribution.

Solution: After applying Bayes' formula (conjugate Gaussian model), the posterior distribution is a Gaussian law $\mathcal{N}(\mu_X, v_X)$, with:

$$\mu_X = \frac{\mu\tau^{-2} + n\sigma^{-2}\bar{X}}{\tau^{-2} + n\sigma^{-2}}, \quad v_X = \frac{1}{\tau^{-2} + n\sigma^{-2}}.$$

ii. We want to test $H_0 = \{\theta \geq 1\}$ against $H_1 = \{\theta < 1\}$ from a Bayesian perspective. Construct the corresponding Bayesian test for the prior Π.

Solution: For the balanced loss function $\ell(\theta, \varphi) = 1_{\theta\in\Theta_0}1_{\varphi=1} + 1_{\theta\in\Theta_1}1_{\varphi=0}$, the optimal Bayesian decision rule is to choose the hypothesis with the highest posterior probability.
We reject H_0 (i.e., $\varphi(X) = 1$) if $\pi(\theta \in H_0|X) \leq \pi(\theta \in H_1|X)$, which is equivalent

to $\pi(H_0|X) \leq 1/2$.

$$\varphi(X) = \mathbb{1}_{\pi(\Theta_0|X)\leq 1/2} = \mathbb{1}_{\pi(\theta\geq 1|X)\leq 1/2}.$$

Since the posterior is a symmetric Gaussian centered at μ_X, the condition $\pi(\theta \geq 1|X) \leq 1/2$ is equivalent to the median (which equals the mean) being less than 1. Thus, the test is: Reject H_0 if $\mu_X < 1$.

iii. What does the test become if we replace H_0 with H_1 and vice-versa?

Solution: If we swap the hypotheses (testing $H_0' : \theta < 1$ vs $H_1' : \theta \geq 1$), the test becomes:

$$\varphi(X) = \mathbb{1}_{\pi(\theta\in(-\infty,1)|X)\leq 1/2}.$$

This corresponds to rejecting H_0' if $\mu_X > 1$.

(b) **Test II:** Let $X = X_1|\theta \sim \mathcal{N}(\theta, 1)$. Consider $H_0^1 : \theta = 0$ vs $H_1^1 : \theta \neq 0$ and $H_0^2 : |\theta| \leq \epsilon$ vs $H_1^2 : |\theta| > \epsilon$.

i. Propose a prior distribution with a Gaussian part $\mathcal{N}(0, \sigma^2)$ for each situation.

Solution:

- **For H_0^1 (Point Null):** Since the null hypothesis is a single point, it is natural (and necessary for the posterior probability of H_0 not to be zero) to choose a prior Π^1 that places a strictly positive mass on $\{0\}$. For example, the "spike-and-slab" prior defined by:

$$\Pi^1 = (1-\alpha)\delta_0 + \alpha\mathcal{N}(0, \sigma^2), \quad \text{with } \alpha \in (0, 1).$$

  This distribution is a mixture of a Dirac mass at 0 and a continuous Gaussian. Note that $\delta_0(\Theta_0) = 1$ and $G(\Theta_1) = 1$.

- **For H_0^2 (Interval Null):** The null hypothesis is composite and forms an interval. A standard diffuse prior is sufficient to assign positive mass to both hypotheses. We can set:

$$\Pi^2 = \mathcal{N}(0, \sigma^2).$$

ii. Compare the corresponding Bayesian tests when ϵ and σ vary.

Solution: In the case of a balanced loss function, the Bayesian test for a prior Π is written as:

$$\varphi(X) = \mathbb{1}_{\pi(\Theta_0|X)\leq\pi(\Theta_1|X)} = \mathbb{1}_{\pi(\Theta_0|X)\leq 1/2}.$$

We determine the posterior laws $\Pi^1[\cdot|X]$ and $\Pi^2[\cdot|X]$ corresponding to the above priors.

The posterior for Π^2 follows the standard conjugate calculation (with $n = 1$):

$$\Pi^2[\cdot|X] = \mathcal{N}\left(\frac{X}{\sigma^{-2}+1}, \frac{1}{\sigma^{-2}+1}\right).$$

The test function $\varphi^2(X)$ for the hypothesis $|\theta| \leq \epsilon$ is:

$$\varphi^2(X) = \mathbb{1}\left\{\Pi^2([-\epsilon, \epsilon] \mid X) \leq \frac{1}{2}\right\}.$$

This compares the mass of the posterior Gaussian falling inside $[-\epsilon, \epsilon]$ to 0.5.
We calculate $\Pi^1(\{0\}|X)$ using Bayes' formula. Let $p_0(X) = \frac{1}{\sqrt{2\pi}}e^{-X^2/2}$ be the likelihood under H_0 ($\theta = 0$), and let $g(\theta)$ be the density of $\mathcal{N}(0, \sigma^2)$.

$$\Pi^1(\{0\}|X) = \frac{(1-\alpha)p_0(X)}{(1-\alpha)p_0(X) + \alpha \int p_\theta(X)g(\theta)d\theta}.$$

We compute the marginal likelihood under the alternative (the integral term):

$$\int p_\theta(X)g(\theta)d\theta = \int \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(X-\theta)^2} \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{\theta^2}{2\sigma^2}} d\theta.$$

This is the convolution of two Gaussians, resulting in $\mathcal{N}(0, 1+\sigma^2)$. The value is:

$$\frac{1}{\sqrt{2\pi(1+\sigma^2)}}e^{-\frac{X^2}{2(1+\sigma^2)}}.$$

Substituting this back, the test $\varphi^1(X)$ rejects H_0 if:

$$\frac{(1-\alpha)e^{-\frac{X^2}{2}}}{(1-\alpha)e^{-\frac{X^2}{2}} + \frac{\alpha}{\sqrt{1+\sigma^2}}e^{-\frac{X^2}{2(1+\sigma^2)}}} \leq \frac{1}{2}.$$

Comparison: One can observe that as $\epsilon \to 0$, the probability assigned to H_0^2 by the continuous posterior goes to 0, leading to rejection of H_0 regardless of X. However, for the point null test, as $\sigma \to \infty$ (vague prior), the denominator term corresponding to H_1 decreases, potentially increasing the probability of H_0. The behaviors are fundamentally different.

4. **Bayes and Constant Risk to Minimax**

(a) Re-prove that a Bayes estimator with constant risk is minimax.

Solution: Let δ_Π be a Bayes estimator for prior Π, so $R_B(\Pi, \delta_\Pi) \leq R_B(\Pi, \delta)$ for any estimator δ. Suppose δ_Π has constant risk $R(\theta, \delta_\Pi) = C$. Then its Bayes risk is also C:

$$R_B(\Pi, \delta_\Pi) = \int R(\theta, \delta_\Pi)d\Pi(\theta) = \int C d\Pi(\theta) = C.$$

Now, let δ^* be any other estimator. Its maximum risk is $\sup_\theta R(\theta, \delta^*)$. We know that the maximum risk is bounded below by the Bayes risk for any prior:

$$\sup_\theta R(\theta, \delta^*) \geq R_B(\Pi, \delta^*) \geq R_B(\Pi, \delta_\Pi) = C.$$

Thus, $\sup_\theta R(\theta, \delta^*) \geq \sup_\theta R(\theta, \delta_\Pi) = C$. Since δ_Π minimizes the maximum risk (it achieves the lower bound C), it is minimax.

(b) Let $X \sim \text{Bin}(n, \theta)$.

i. Show that the family of priors $\{\Pi_{a,b} = \text{Beta}(a, b)\}$ is conjugate.

Solution: Likelihood: $L(\theta) \propto \theta^x(1-\theta)^{n-x}$. Prior: $\pi(\theta) \propto \theta^{a-1}(1-\theta)^{b-1}$. Posterior: $\pi(\theta|x) \propto \theta^{x+a-1}(1-\theta)^{n-x+b-1}$. This is the density of a $\text{Beta}(a+x, b+n-x)$ distribution, up to a constant factor.

ii. Give a Bayes estimator $\hat{\theta}_{a,b}(X)$ for $\Pi_{a,b}$ and the quadratic loss.

Solution: For quadratic loss, the Bayes estimator is the posterior mean:

$$\hat{\theta}_{a,b}(X) = \frac{a + X}{a + b + n}.$$

iii. Assume $a = b$. Find a minimax estimator for the quadratic loss.

Solution: According to the first question, it suffices to find an estimator among the family $\hat{\theta}_{a,b}(X)$ that has constant risk. If such an estimator exists, since it is a Bayes estimator (as established previously), it will be minimax.
The problem suggests setting $a = b$. We therefore calculate the quadratic risk of the estimator $\hat{\theta}_{a,a}(X)$:

$$\begin{aligned}
\mathbb{E}_\theta[(\hat{\theta}_{a,a}(X) - \theta)^2] &= \mathbb{E}_\theta\left[\left(\frac{(X - n\theta) + a(1 - 2\theta)}{2a + n}\right)^2\right] \\
&= \frac{\text{Var}_\theta(X)}{(2a + n)^2} + \frac{a^2(1 - 2\theta)^2}{(2a + n)^2} \\
&= (2a + n)^{-2}\left[n\theta(1 - \theta) + a^2(1 - 2\theta)^2\right] \\
&= (2a + n)^{-2}\left[(4a^2 - n)\theta^2 + (n - 4a^2)\theta + a^2\right].
\end{aligned}$$

For the risk to be constant (independent of θ), the coefficients involving θ must vanish. We conclude that if we choose $n = 4a^2$, which implies:

$$a = a_n = \frac{\sqrt{n}}{2},$$

the estimator $\hat{\theta}_{a_n,a_n}(X)$ has constant risk. Being a constant risk Bayes estimator, it is minimax.

iv. Is the estimator $T = X/n$ minimax?

Solution: To determine if T is minimax, it suffices to compare its maximal risk $R_{\max}(T)$ to the minimax risk R_M. We can calculate R_M because we found a minimax estimator in the previous question, $\hat{\theta}_{a_n,a_n}(X)$. The maximal risk of this estimator corresponds to its constant risk:

$$R_{\max}(\hat{\theta}_{a_n,a_n}(X)) = \frac{a_n^2}{(2a_n+n)^2} = \frac{n}{4(n+\sqrt{n})^2}.$$

Now, we calculate the maximal risk of $T = X/n$:

$$\begin{aligned} R_{\max}(T) &= \sup_{\theta\in(0,1)} R(\theta, T) \\ &= \sup_{\theta\in(0,1)} \frac{\mathrm{Var}_\theta(X)}{n^2} \\ &= \sup_{\theta\in(0,1)} \frac{n\theta(1-\theta)}{n^2} = \frac{1}{4n}, \end{aligned}$$

since the function $\theta \mapsto \theta(1-\theta)$ reaches its maximum at $\theta = 1/2$, where the value is $1/4$.
Comparing the two risks:

$$R_{\max}(T) = \frac{1}{4n} > \frac{1}{4(\sqrt{n}+1)^2} = R_M.$$

We conclude that $R_{\max}(T) > R_M$, and therefore T is **not** minimax.

5. **Bayes and Unique to Admissible**

(a) Give a Bayes estimator for Π. We will denote it T_1.

Solution: For quadratic loss, $T_1(x) = \mathbb{E}[\theta|X=x] = \int \theta d\Pi(\theta|x)$.

(b) Let $m^\pi(x) = \int f_\theta(x)d\Pi(\theta)$. How is this quantity interpreted?

Solution: $m^\pi(x)$ is the marginal density of the observation X under the prior Π. It represents the weighted average of the likelihoods.

(c) Show that $R_B(\Pi, T) = \int \mathbb{E}[(T(X)-\theta)^2|X=x]m^\pi(x)d\mu(x)$.

Solution: By definition, the Bayes risk is the joint expectation of the loss:

$$R_B(\Pi, T) = \iint (T(x)-\theta)^2 f_\theta(x)d\mu(x)d\Pi(\theta).$$

Using Fubini's theorem to swap integrals and the definition of the posterior $d\Pi(\theta|x) = \frac{f_\theta(x)d\Pi(\theta)}{m^\pi(x)}$, we write $f_\theta(x)d\Pi(\theta) = m^\pi(x)d\Pi(\theta|x)$.

$$R_B(\Pi, T) = \int \left(\int (T(x)-\theta)^2 d\Pi(\theta|x)\right) m^\pi(x)d\mu(x).$$

The inner integral is exactly $\mathbb{E}[(T(X)-\theta)^2|X=x]$.

(d) Show that if $dQ = m^\pi d\mu$ dominates all P_θ, then equivalent Bayes estimators have the same risk everywhere.

Solution: If T_1 and T_2 are both Bayes estimators, they minimize the inner integral $\mathbb{E}[(\cdot-\theta)^2|X=x]$ for m^π-almost all x. Since the minimizer (posterior mean) is unique, $T_1(x) = T_2(x)$ almost everywhere with respect to the marginal measure Q. If Q dominates all P_θ, then any set of measure zero under Q also has measure zero under P_θ for all θ. Thus $T_1 = T_2$ P_θ-a.s., implying their risks $R(\theta, T_1) = \mathbb{E}_\theta[(T_1-\theta)^2]$ and $R(\theta, T_2)$ are identical.

(e) Show that if the Bayes estimator is unique up to equivalence, it is admissible.

Solution: Let T be the unique Bayes estimator. Suppose T is not admissible. Then there exists T' such that $R(\theta, T') \leq R(\theta, T)$ for all θ, with strict inequality for some θ_0. Taking the expectation over the prior Π:

$$R_B(\Pi, T') = \int R(\theta, T')d\Pi(\theta) \leq \int R(\theta, T)d\Pi(\theta) = R_B(\Pi, T).$$

Since T is Bayes, $R_B(\Pi, T) \leq R_B(\Pi, T')$. Therefore, $R_B(\Pi, T') = R_B(\Pi, T)$, so T' is also a Bayes estimator. Since the Bayes estimator is unique up to equivalence, T' and T have the same frequentist risk. This contradicts the assumption that T' strictly dominates T at θ_0. Thus, T must be admissible.

(f) **Application:** Let $\mathcal{P} = \{P_\theta = \mathcal{N}(\theta, 1), \theta \in \mathbb{R}\}$ and $X_1, \ldots, X_n$ i.i.d. with law P_θ given θ. Set $\Pi = \mathcal{N}(a, \sigma^2)$, with $a \in \mathbb{R}$ and $\sigma^2 > 0$ fixed.

i. Calculate the Bayes estimator for the prior Π and quadratic loss.

Solution: The posterior mean is Bayes and is equal here to:

$$\hat{\theta}_{a,\sigma^2} = \frac{n\overline{X} + a}{n + \sigma^{-2}}.$$

ii. Determine the marginal distribution of $X = (X_1, \ldots, X_n)$, denoted Q_n.

Solution: We note that $X' = (X_1, \ldots, X_n)^T$ has the same distribution as the column vector Y whose coordinates are given by

$$Y_i = \theta + \epsilon_i,$$

where the ϵ_i are i.i.d. $\mathcal{N}(0, 1)$ and θ is a random variable independent of the ϵ_i and distributed as $\mathcal{N}(a, \sigma^2)$. Equivalently,

$$Y_i = a + \sigma\zeta + \epsilon_i,$$

with $\zeta \sim \mathcal{N}(0,1)$ independent of the ϵ_i. The vector Y is Gaussian, with mean $a\mathbf{1}$ (where $\mathbf{1}$ is the column vector of ones) and variance-covariance matrix:

$$\mathrm{Var}(Y) = \begin{bmatrix} 1+\sigma^2 & \sigma^2 & \dots & \sigma^2 \\ \sigma^2 & 1+\sigma^2 & \dots & \sigma^2 \\ \vdots & & \ddots & \vdots \\ \sigma^2 & \dots & \dots & 1+\sigma^2 \end{bmatrix} = I_{n\times n} + \sigma^2 \mathbf{1}\mathbf{1}^T.$$

We deduce that $X \sim \mathcal{N}(a\mathbf{1}, \Sigma)$, with $\Sigma = I_{n\times n} + \sigma^2\mathbf{1}\mathbf{1}^T$. Note that Σ is invertible, for example because it is positive definite since $y^T\Sigma y = \|y\|_2^2 + \sigma^2(\sum y_i)^2 > 0$ as soon as $y \neq 0$.

iii. Verify that Q_n dominates all laws $P_\theta^{\otimes n}$.

Solution: The distribution Q_n has a density with respect to the Lebesgue measure on $\mathbb{R}^n$ given by the standard Gaussian formula. This density is strictly positive on all of $\mathbb{R}^n$. It follows that if $Q_n(A) = 0$ for a measurable set A, then A is of Lebesgue measure zero on $\mathbb{R}^n$, and thus $P_\theta^{\otimes n}(A) = 0$ for all θ. Thus Q_n dominates all laws $P_\theta^{\otimes n}$.

iv. Show that the estimators $\alpha\overline{X} + \beta$, with $\alpha \in [0,1)$ and $\beta \in \mathbb{R}$, are admissible.

Solution: Note that a Bayes estimator for the prior $\Pi = \mathcal{N}(a, \sigma^2)$ is written, according to part i., as $\alpha\overline{X} + \beta$, with $\alpha = \frac{n}{n+\sigma^{-2}}$ and $\beta = \frac{a}{n+\sigma^{-2}}$. Moreover, from part iii. and the first part of the exercise (dominance implies uniqueness up to equivalence), it is unique up to equivalence and therefore admissible.
By varying a in $\mathbb{R}$ and σ^2 in $(0, +\infty)$, we obtain all estimators $\alpha\overline{X} + \beta$ with $\alpha \in (0,1)$ and $\beta \in \mathbb{R}$.
The case $\alpha = 0$ is obtained otherwise by noting that the constant estimator equal to β is admissible. Indeed, it it were inadmissible, there would exist another estimator T whose frequentist risk satisfies $R(\beta, T) \leq R(\beta, \beta) = 0$. This implies that T is equal to β P_β-almost surely. Since P_β is equivalent to any other P_θ, we deduce that T is equal to β P_θ-almost surely and $R(\theta, T) \leq R(\theta, \beta)$ for any θ, contradicting inadmissibility.

v. Show that the estimators $\overline{X} + \beta$, for $\beta \neq 0$, are not admissible.

Solution: The estimators $\overline{X} + \beta$, for $\beta \neq 0$, are not admissible, because their quadratic risk is strictly greater than that of the estimator $\overline{X}$:

$$R(\theta, \overline{X} + \beta) = \mathbb{E}_\theta[(\overline{X} - \theta + \beta)^2] = \frac{1}{n} + \beta^2 > \frac{1}{n} = R(\theta, \overline{X}).$$