# 1: Hypothesis testing with multiple Gaussian observations

Let $\sigma_1, \ldots, \sigma_n > 0$ be strictly positive numbers (which we assume to be known), and define $\Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_n^2)$.

Consider the statistical model $P_\theta = N(\theta, \Sigma)$, indexed by $\theta \in \Theta = \mathbb{R}$. In other words, $P_\theta$ is the joint distribution of $n$ independent Gaussian observations $X_1, \ldots, X_n$, where $X_i \sim N(\theta, \sigma_i^2)$.

(a) Consider testing $H_0 : \theta \leq 0$ versus $H_1 : \theta > 0$ at level $\alpha \in (0, 1)$. Is there a uniformly most powerful (UMP) test? If yes, construct the test and prove that it is UMP. If not, prove your claim.

(b) Consider next $\tilde{H}_0 : \theta = 0$ versus $\tilde{H}_1 : \theta \neq 0$ at level $\alpha \in (0, 1)$. Is there a uniformly most powerful (UMP) test? If yes, construct the test and prove that it is UMP. If not, prove your claim.

# 2: Lower bounds

We consider $n$ i.i.d. samples $X_1, \ldots, X_n \sim_{\text{iid}} P_\theta^{(1)}$ where $P_\theta^{(1)}$ is the distribution with density

$$p_\theta^{(1)}(x) = 1_{x \geq \theta} e^{-(x-\theta)}, \quad (1)$$

In other words, $X_i = \theta + W_i$, where $(W_i)_{i \leq n}$ are i.i.d. exponential random variables with rate one. We let $P_\theta$ be the law of $X = (X_1, \ldots, X_n)$ (or equivalently $P_\theta = (P_\theta^{(1)})^{\otimes n}$), with $\theta \in \Theta = \mathbb{R}$. We will use the square loss $L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$.

(a) Let $X_{(1)} < X_{(2)} \leq \cdots \leq X_{(n)}$ be the order statistics of $X$ (notice that the coordinates of $X$ are almost surely distinct). Consider the estimator $\hat{\theta}(X) = X_{(1)} - (1/n)$. Compute its risk function $R(\hat{\theta}; \theta)$, and its worst case risk $R_M(\hat{\theta}; \Theta) = \sup_{\theta \in \Theta} R(\hat{\theta}; \theta)$.

(b) Does the Crâmer-Rao lower bound apply to this case? Justify your answer.

(c) Use Le Cam’s method to show that that the estimator $\hat{\theta}$ defined above is optimal up to constants. Namely, prove that there exists a numerical constant $C$ such that the minimax risk satisfies $R_M(\Theta) \leq R_M(\hat{\theta}; \Theta) \leq C R_M(\Theta)$. 

1
(d) Is the estimator $\hat{\theta}$ exactly minimax optimal? Justify your answer.

(e) We now generalize the above model. Namely, let $W_1, \ldots, W_n$ be i.i.d. with common density $q$ on $\mathbb{R}$. We observe $X_i = \theta + W_i$, $i \in \{1, \ldots, n\}$. The parameter space is, as above, $\Theta = \mathbb{R}$.

Fix $\eta > 0$. Can you find $q$ so that (for this new problem) $c_1 n^{-1-\eta} \leq R_M(\Theta) \leq c_2 n^{-1-\eta}$ for two constants $c_1, c_2 > 0$? Prove your answer.

# 3: More Gaussian hypothesis testing

Consider the Gaussian mean model $P_\theta = N(\theta, 1)$, $\theta \in \Theta = \mathbb{R}^d$, $d \geq 2$. Is there a uniformly most powerful unbiased (UMPU) test for testing $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$. Justify your answer. **Remark:** At level $\alpha \in (0,1)$

Maybe have a part a) where they need to do $d = 1$?

**Solution:**

Let us first start with the 1-dimensional problem, i.e. we observe $Z \sim N(\theta, 1)$ and want to test $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$. This is a one-parameter exponential family, and the results from Section 4.2. of TSH are immediately applicable to derive a UMPU test. In particular, a level $\alpha$ UMPU test has the following form:

$$\phi(Z) = \begin{cases} 1 & |Z| \geq z_{1-\alpha/2} \\ 0 & |Z| < z_{1-\alpha/2} \end{cases}$$

Here $z_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard Normal distribution. Furthermore, up to Lebesgue-measure zero, this is the unique UMPU test (arguing as in Theorem 3.6.1 in TSH).

Moving to the $d$-dimensional problem, let us assume there does exist a UMPU test, let us call it $\psi(X)$, where $X = (X_1, \ldots, X_d) \sim N(\theta, 1)$ (as seen in Homework 7, we may assume the test is a function of $X$).

Let us also consider the following tests for $j \in \{1, 2\}$:

$$\phi_j(X) = \begin{cases} 1 & |X_j| \geq z_{1-\alpha/2} \\ 0 & |X_j| < z_{1-\alpha/2} \end{cases}$$

Let us note these tests are unbiased, since e.g. for any $\theta$:

$$E_{\theta}[\phi_j(X)] = P_{\theta}[|X_j| \geq z_{1-\alpha/2}] = P_{\theta_j}[|X_j| \geq z_{1-\alpha/2}] \geq \alpha$$

The last step follows from direct calculation (taking the derivative w.r.t. $\theta_j$ we see that it is minimized at $\theta_j = 0$ at which it equals $\alpha$), but is also a direct consequence of the fact that $\phi$ is unbiased in the 1-dimensional problem.

Now, the test $\psi$ must be at least as powerful at $\phi_1$ and $\phi_2$, since it is UMPU and the other two tests are unbiased.

However, let us now restrict ourselves to the subfamily of alternatives consisting of $\theta = (\theta_1, 0, \ldots, 0), \theta_1 \neq 0$. Under these alternatives, $X_1$ is sufficient, hence again applying the result of Homework 7, and applying our 1-dimensional result, we see that the test $\phi(X_1)$ must be UMPU. This means that $\phi_1(X)$ is UMPU and (repeating the argument from the 1D case) we see that it also is uniquely so (up to Lebesgue measure zero). But $\psi(X)$ is at least as powerful as $\phi_1(X)$, since both are UMPU for the full problem, and thus must be UMPU for the subproblem too. By uniqueness, we see that this implies that $\psi(X)$ must reject when $|X_1| > 1 - z_{1-\alpha/2}$ (up to Lebesgue-measure zero sets).

Now reapplying the same argument to the subfamily of alternatives of $\theta$, where only the 2nd coordinates is nonzero, we see that $\psi$ must also reject when $|X_2| > 1 - z_{1-\alpha/2}$.

But this is a contradiction, since then $\psi$ is not a level $\alpha$ test, i.e. $E_{\theta}[\psi(X)] > \alpha$. Thus no UMPU test exists for $d \geq 2$. 

2
### 4: Bowl-shaped losted

Let $q$ be a probability density function on $\mathbb{R}$, with zero mean and finite moments of all orders. We observe $X = \theta + W$ for $\theta \in \mathbb{R}$ and $W \sim q$. We will consider a loss function $L(\hat{\theta}, \theta) = \ell(\hat{\theta} - \theta)$ where $\ell$ is bowl-shaped (in this case, this means that $\ell(-t) = \ell(t)$ and $t \mapsto \ell(t)$ is non-decreasing for $x \geq 0$).

(a) Let $\hat{\theta}(x) = x$. Show that this is minimax optimal for the square loss $\ell(t) = t^2$.

(b) Construct a case (i.e. a density $q$, and a loss $\ell$) for which $\hat{\theta}(x) = x$ is not minimax optimal. Justify your conclusion.

Remark: For (b) $q$ and $\ell$ have to satisfy properties specified above?

Solution:

(a) First, $\hat{\theta}(x)$ has constant risk, namely for all $\theta \in \mathbb{R}$ (since the distribution of $W$ is $q$ which does not depend on $\theta$):

$$E_\theta[(\hat{\theta}(X) - \theta)^2] = E_\theta[(W + \theta - \theta)^2] = E[W^2]$$

For $\theta$, we will consider the sequence of uniform priors $\pi_m = U[-m, m]$ for $m \in \mathbb{N}$. Let us call $\delta_m(X) = E_{\pi_m}[\theta \mid X]$ the Bayes estimator for $\theta$ and $r_m = E_{\pi_m}[(\theta - \delta_m(X))^2]$ the Bayes risk with respect to the prior $\pi_m$.

Note that over all estimator $\hat{\theta}(X)$ of $\theta$ we may restrict our attention to non-randomized estimators that are a function of $X$, since the square loss is convex and hence we may always apply Rao-Blackwell, it holds that:

$$\inf_{\hat{\theta}} \sup_{\theta \in \mathbb{R}} E_\theta[(\hat{\theta}(X) - \theta)^2] \geq r_m$$

Below we will argue that $r_m \to E[W^2]$ as $m \to \infty$. Then taking $m \to \infty$, we see that:

$$\inf_{\hat{\theta}} \sup_{\theta \in \mathbb{R}} E_\theta[(\hat{\theta}(X) - \theta)^2] \geq E[W^2] = \sup_{\theta \in \mathbb{R}} E_\theta[(\hat{\theta}(X) - \theta)^2]$$

This implies that $\hat{\theta}(X) = X$ is indeed minimax optimal. Let us now prove our claim for $r_m$. Since $r_m \leq \sup_{\theta \in \mathbb{R}} E_\theta[(\hat{\theta}(X) - \theta)^2] = E[W^2]$, it suffices to argue that

$$\lim_{m \to \infty} r_m \geq E[W^2]$$

For some intuition, let us study:

$$\delta_m(x) = E_{\pi_m}[\theta \mid X = x] = \frac{\int_{-m}^{m} uq(x-u)du}{\int_{-m}^{m} q(x-u)du}$$

Note that since $q$ is a density, as $m \to \infty$, we have for almost every $x$ (w.r.t. Lebesgue measure) that $\int_{-m}^{m} q(x-u)du \to 1$. Furthermore, since for almost every $x$ it holds that $\int_{-\infty}^{\infty} |u|q(x-u)du < \infty$, we similarly get from dominated convergence that for almost every $x$:

$$\lim_{m \to \infty} \int_{-m}^{m} uq(x-u)du = \int_{-\infty}^{\infty} uq(x-u)du = x$$
Here we also used that $\int_{-\infty}^{\infty} \theta q(\theta) d\theta = E[W] = 0$.

The above imply that $\delta_m(x) = E_{\pi_m}[\theta \mid X = x] \to x$ almost everywhere as $m \to \infty$. This provides a hint that $\theta(X) = X$ is minimax optimal, but we are not quite done yet.

Let us look once more at $\delta_m(x)$, where we now decompose $x = w + \theta$. Upon substituting $\tilde{\theta}$ for $u - \theta$ we see that:

$$\delta_m(w + \theta) = \frac{\int_m \theta q(w + \theta - u) du}{\int_m \theta q(w + \theta - u) du + \int_m \tilde{\theta} q(w - \tilde{\theta}) d\tilde{\theta}}$$

Now we turn to $r_m$, where now we let $Z \sim U[-1,1]$ so that $mZ \sim \pi_m$:

$$r_m^2 = E_{\pi_m}[(\delta_m(X) - \theta)^2]$$

$$= E_{\pi_m}[(\delta_m(W + \theta) - \theta)^2]$$

$$= E_{\pi_m} \left[ \left( \frac{\int_m \theta q(w - u) du}{\int_m \theta q(w - u) du + \int_m \tilde{\theta} q(w - \tilde{\theta}) d\tilde{\theta}} \right)^2 \right]$$

Note that the last expression is an expectation over the independent $Z \sim U[-1,1]$ and $W \sim q$. Now arguing as we did for the asymptotics of $\delta_m(x)$, we see that almost surely:

$$\left( \frac{\int_m \theta q(w - u) du}{\int_m \theta q(w - u) du + \int_m \tilde{\theta} q(w - \tilde{\theta}) d\tilde{\theta}} \right)^2 \to W^2 \text{ as } m \to \infty$$

(Since for example $m(1 - Z) \to \infty$ almost surely, noting that $Z < 1$ with probability 1.)

Hence applying Fatou’s Lemma we see that:

$$\liminf_{m \to \infty} r_m \geq E \left[ \liminf_{m \to \infty} \left( \frac{\int_m \theta q(w - u) du}{\int_m \theta q(w - u) du + \int_m \tilde{\theta} q(w - \tilde{\theta}) d\tilde{\theta}} \right)^2 \right] = E[W^2]$$

Thus we conclude.

(b) Let us take $\ell(t) = 1_{\{|t| > 1\}}$, which is Bowl-shaped and further let:

$$q(w) = \frac{1}{2} 1_{\{w \in (-2,-1) \cup (1,2)\}}$$

$q$ satisfies the required assumptions. Next, note that $|W| > 1$ almost surely, so that for any $\theta$, $\ell(\hat{\theta}(X) - \theta) = \ell(\theta + W - \theta) = \ell(W) = 1$ almost surely. Thus:

$$\sup_{\theta \in \mathbb{R}} E_{\theta}[L(\hat{\theta}(X), \theta)] = 1$$

On the other hand, consider the estimator:
\[ \hat{\delta}(X) = X - 1 \]

Then note that:
\[ \ell(\hat{\delta}(X) - \theta) = \ell(W - 1) = 1_{\{|W-1| > 1\}} \]

But since \( \Pr[|W - 1| > 1] = 1/2 \), we get:
\[ \sup_{\theta \in \mathbb{R}} E_\theta[L(\hat{\delta}(X), \theta)] = \frac{1}{2} \]

And therefore, \( \hat{\theta}(X) \) cannot be minimax optimal.