# Homework 4

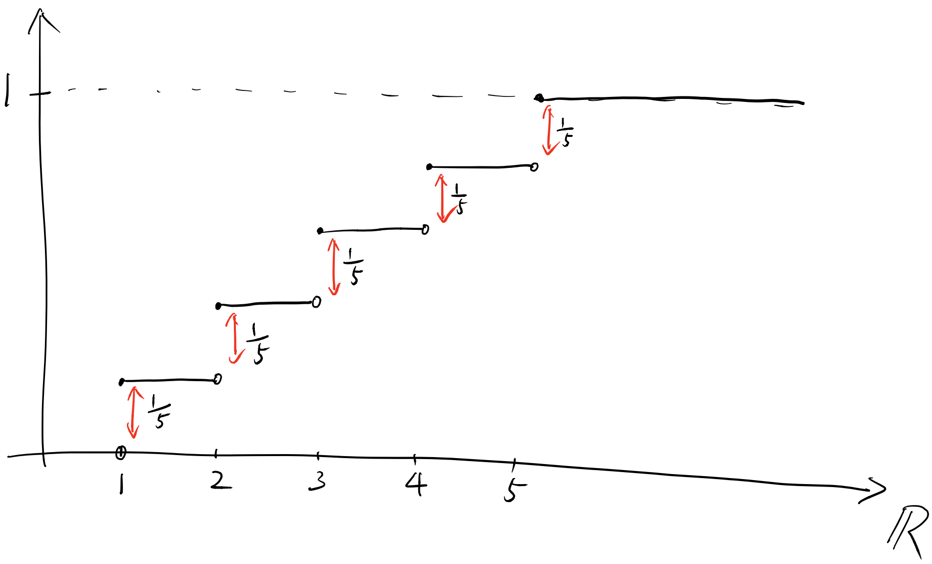
**Problem 1.** For an i.i.d. random sample $\{X_1, ..., X_n\}$ , suppose that the CDF of $X_i$ is $F_X(x) = \Pr(X_i \leq x)$. Define the empirical CDF $\widehat{F}_X(x) = \frac{1}{n}\sum_{i=1}^n 1(X_i \leq x)$. For any fixed $x \in \mathbb{R}$, show that $\widehat{F}_X(x)$ is a consistent and unbiased estimator of $F_X(x)$. Suppose $n = 5$, $X_1 = 1$, $X_2 = 2$, $X_3 = 3$, $X_4 = 4$ and $X_5 = 5$. Can you draw the graph of the function $\widehat{F}_X : \mathbb{R} \to [0, 1]$?

**Solution.** Unbiasedness:

$$\mathbb{E}\left(\widehat{F}_X(x)\right) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}1(X_i \leq x) = \frac{1}{n}\sum_{i=1}^n (1 \cdot \Pr(X_i \leq x) + 0 \cdot \Pr(X_i > x)) = F_X(x).$$

Consistency: by WLLN,

$$\frac{1}{n}\sum_{i=1}^n 1(X_i \leq x) \to_p \mathbb{E}1(X_i \leq x) = F_X(x).$$




**Problem 2.** For an i.i.d. random sample $\{X_1, ..., X_n\}$ with $X_i \sim N(\mu, 1)$. We are interested in estimating $\mu$. We know that $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \sim N\left(\mu, \frac{1}{n}\right)$. Suppose $\mu$ is realization of some random variable $M$ distributed as $N\left(\alpha, \beta^2\right)$. $\alpha$ and $\beta$ are fixed constants. $\overline{X}_n \sim N\left(\mu, \frac{1}{n}\right)$ should be interpreted as "the conditional distribution of $\overline{X}$ given $M = \mu$ is $\overline{X}_n \sim N\left(\mu, \frac{1}{n}\right)$". Recall the Bayes theorem: for two random variables $X, Y$,

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x) f_X(x)}{\int f_{Y|X}(y|x) f_X(x) \, \mathrm{d}x}.$$

Now you can apply Bayes theorem to derive the conditional density of $M$ given $\overline{X}_n$.

(i) Show that the marginal distribution of $\overline{X}_n$ is $N\left(\alpha, \frac{1}{n} + \beta^2\right)$. Hint: Show that

$$n\left(\overline{X}_n - \mu\right)^2 + \frac{(\mu-\alpha)^2}{\beta^2} - \frac{\left(\overline{X}_n - \alpha\right)^2}{\frac{1}{n} + \beta^2} = \frac{\left(n\beta^2\left(\overline{X}_n - \mu\right) - (\mu - \alpha)\right)^2}{(n\beta^2+1)\,\beta^2}.$$

(ii) Define the Bayes estimator $\widehat{\mu}_{\alpha,\beta}$ (which is indeed of great importance in statistics) to be

$$\widehat{\mu}_{\alpha,\beta} = \int z f_{M|\overline{X}_n}\left(z|\overline{X}_n\right) \mathrm{d}z,$$

the mean of the conditional distribution of $M$ given $\overline{X}_n$. Show that

$$\widehat{\mu}_{\alpha,\beta} = \frac{\beta^2}{\beta^2 + \frac{1}{n}}\overline{X}_n + \frac{\frac{1}{n}}{\beta^2+\frac{1}{n}}\alpha.$$

Hint: Use the Bayes theorem to show that the conditional distribution of $M$ given $\overline{X}_n$ is

$$N\left(\frac{\beta^2}{\beta^2 + \frac{1}{n}}\overline{X}_n + \frac{\frac{1}{n}}{\beta^2+\frac{1}{n}}\alpha, \frac{\frac{1}{n}\beta^2}{\frac{1}{n}+\beta^2}\right).$$

(iii) Is $\widehat{\mu}_{\alpha,\beta}$ unbiased?

(iv) Is $\widehat{\mu}_{\alpha,\beta}$ consistent?

**Solution.** The joint density of $\left(\overline{X}_n, M\right)$ is

$$f_{\overline{X}_n,M}(x,\mu) = \frac{1}{\sqrt{2\pi\frac{1}{n}}}\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\frac{1}{\sqrt{n}}}\right)^2\right)\frac{1}{\sqrt{2\pi\beta^2}}\exp\left(-\frac{1}{2}\left(\frac{\mu-\alpha}{\beta}\right)^2\right).$$

The marginal density of $\overline{X}_n$ is

$$\begin{aligned}
\int f_{\overline{X}_n,M}(x,\mu)\,\mathrm{d}\mu &= \int \frac{1}{\sqrt{2\pi\frac{1}{n}}}\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\frac{1}{\sqrt{n}}}\right)^2\right)\frac{1}{\sqrt{2\pi\beta^2}}\exp\left(-\frac{1}{2}\left(\frac{\mu-\alpha}{\beta}\right)^2\right)\mathrm{d}\mu \\
&= \exp\left(-\frac{1}{2}\frac{(x-\alpha)^2}{\frac{1}{n}+\beta^2}\right)\frac{1}{\sqrt{2\pi\frac{1}{n}}}\frac{1}{\sqrt{2\pi\beta^2}}\int \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\frac{1}{\sqrt{n}}}\right)^2 - \frac{1}{2}\left(\frac{\mu-\alpha}{\beta}\right)^2 + \frac{1}{2}\frac{(x-\alpha)^2}{\frac{1}{n}+\beta^2}\right)\mathrm{d}\mu \\
&= \exp\left(-\frac{1}{2}\frac{(x-\alpha)^2}{\frac{1}{n}+\beta^2}\right)\frac{1}{\sqrt{2\pi\frac{1}{n}}}\frac{1}{\sqrt{2\pi\beta^2}}\int \exp\left(-\frac{1}{2}\frac{\left(n\beta^2(x-\mu)-(\mu-\alpha)\right)^2}{(n\beta^2+1)\,\beta^2}\right)\mathrm{d}\mu \\
&= \exp\left(-\frac{1}{2}\frac{(x-\alpha)^2}{\frac{1}{n}+\beta^2}\right)\frac{\sqrt{2\pi\cdot\frac{\beta^2}{n\beta^2+1}}}{\sqrt{2\pi\frac{1}{n}}\sqrt{2\pi\beta^2}}\frac{1}{\sqrt{2\pi\cdot\frac{\beta^2}{n\beta^2+1}}}\int \exp\left(-\frac{1}{2}\frac{\left(\mu-\frac{\alpha+\beta^2 nx}{n\beta^2+1}\right)^2}{\frac{\beta^2}{n\beta^2+1}}\right)\mathrm{d}\mu \\
&= \frac{1}{\sqrt{2\pi\frac{n\beta^2+1}{n}}}\exp\left(-\frac{1}{2}\frac{(x-\alpha)^2}{\frac{1}{n}+\beta^2}\right).
\end{aligned}$$

The conditional density of $M$ given $\overline{X}_n$:

$$f_{M|\overline{X}_n}(\mu|x) = \frac{f_{\overline{X}_n,M}(x,\mu)}{f_{\overline{X}_n}(x)} = \frac{\frac{1}{\sqrt{2\pi\frac{1}{n}}}\exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\frac{1}{\sqrt{n}}}\right)^2\right)\frac{1}{\sqrt{2\pi\beta^2}}\exp\left(-\frac{1}{2}\left(\frac{\mu-\alpha}{\beta}\right)^2\right)}{\frac{1}{\sqrt{2\pi\frac{n\beta^2+1}{n}}}\exp\left(-\frac{1}{2}\frac{(x-\alpha)^2}{\frac{1}{n}+\beta^2}\right)}$$

$$= \frac{\sqrt{2\pi \frac{n\beta^2+1}{n}}}{\sqrt{2\pi \frac{1}{n}}\sqrt{2\pi\beta^2}} \exp\left( -\frac{1}{2}\left(\frac{x-\mu}{\frac{1}{\sqrt{n}}}\right)^2 - \frac{1}{2}\left(\frac{\mu-\alpha}{\beta}\right)^2 + \frac{1}{2}\frac{(x-\alpha)^2}{\frac{1}{n}+\beta^2} \right)$$

$$= \frac{1}{\sqrt{2\pi \cdot \frac{\beta^2}{n\beta^2+1}}} \exp\left( -\frac{1}{2}\frac{\left(\mu - \frac{\alpha+\beta^2 nx}{n\beta^2+1}\right)^2}{\frac{\beta^2}{n\beta^2+1}} \right).$$

$\widehat{\mu}_{\alpha,\beta}$ is not unbiased:

$$\mathbb{E}\widehat{\mu}_{\alpha,\beta} = \frac{\beta^2}{\beta^2+\frac{1}{n}}\mu + \frac{\frac{1}{n}}{\beta^2+\frac{1}{n}}\alpha \neq \mu.$$

$\widehat{\mu}_{\alpha,\beta}$ is consistent: by WLLN and continuous mapping theorem,

$$\widehat{\mu}_{\alpha,\beta} = \frac{\beta^2}{\beta^2+\frac{1}{n}}\overline{X}_n + \frac{\frac{1}{n}}{\beta^2+\frac{1}{n}}\alpha \to_p \frac{\beta^2}{\beta^2+0}\mu + \frac{0}{\beta^2+0}\alpha = \mu.$$

**Problem 3.** Consider the following two regression models.

**Model 1**: $Y_i = \boldsymbol{X}_i'\boldsymbol{\gamma}_1 + e_i$ if $i = 1, \ldots, n_1$, and $Y_i = \boldsymbol{X}_i'\boldsymbol{\gamma}_2 + u_i$ if $i = n_1+1, \ldots, n$, where $\boldsymbol{X}_i$ is the $k$-vector of regressors, $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ are the unknown $k$-vectors of parameters. This means we split the sample into two and run two regressions.

**Model 2**: $Y_i = \boldsymbol{X}_i'\boldsymbol{\gamma}_1 + (d_i\boldsymbol{X}_i)'\,\boldsymbol{\delta} + v_i$, $i = 1, \ldots, n$, where $\boldsymbol{\gamma}_1$ and $\boldsymbol{\delta}$ are the unknown $k$-vectors of parameters, and $d_i$ is the dummy variable:

$$d_i = \begin{cases} 0 & \text{for } i = 1, \ldots n_1, \\ 1 & \text{for } i = n_1+1, \ldots, n. \end{cases}$$

Show that the two regression models give the same fitted values and residuals by following the steps below.

**(a)** Show that Models 1 and 2 can be written as $\boldsymbol{Y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{e}$ and $\boldsymbol{Y} = \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{v}$ respectively for some $n \times 2k$ matrices $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, and $2k$-vectors $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ (describe the matrix $\boldsymbol{X}$ and the vector $\boldsymbol{\beta}$ in each model).

**(b)** Show that $\boldsymbol{X}_2 = \boldsymbol{X}_1\boldsymbol{A}$ (find $\boldsymbol{A}$).

**(c)** Show that $\boldsymbol{A}$ is invertible.

**(d)** Show that the two regression models give the same fitted values and residuals.

**Solution.** Let

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{X}_1' \\ \boldsymbol{X}_2' \\ \vdots \\ \boldsymbol{X}_{n_1}' \end{bmatrix}$$

be a $n_1 \times k$ matrix and let

$$\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{X}_{n_1+1}' \\ \boldsymbol{X}_{n_1+2} \\ \vdots \\ \boldsymbol{X}_n' \end{bmatrix}$$

be a $(n-n_1) \times k$ matrix. Define the $n \times 2k$ matrix $\boldsymbol{X}_1 = \begin{bmatrix} \boldsymbol{W} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Z} \end{bmatrix}$ and let $\boldsymbol{\beta}_1 = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}$ be a $2k \times 1$ vector. Finally, define the $n \times 2k$ matrix $\boldsymbol{X}_2 = \begin{bmatrix} \boldsymbol{W} & \boldsymbol{0} \\ \boldsymbol{Z} & \boldsymbol{Z} \end{bmatrix}$ and let $\boldsymbol{\beta}_2 = \begin{bmatrix} \gamma_1 \\ \boldsymbol{\delta} \end{bmatrix}$ be a $2k \times 1$ vector of parameters. a)

Using the above definitions, it is easy to verify that Model 1 can be written as $\boldsymbol{Y} = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{e}$ and Model 2 can be written as $\boldsymbol{Y} = \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{v}$, where $\boldsymbol{Y}$ is a $n \times 1$ vector of observations and $\boldsymbol{e}$ and $\boldsymbol{v}$ are $n \times 1$ vectors of errors. b) What we need to find is a $2k \times 2k$ matrix $\boldsymbol{A}$ such that $\boldsymbol{X}_2 = \boldsymbol{X}_1\boldsymbol{A}$. Let $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 \\ \boldsymbol{A}_3 & \boldsymbol{A}_4 \end{bmatrix}$ where $\boldsymbol{A}_i$ is a $k \times k$ matrix. Then, $\boldsymbol{X}_2 = \boldsymbol{X}_1\boldsymbol{A}$ implies that $\boldsymbol{W}\boldsymbol{A}_1 = \boldsymbol{W}$; $\boldsymbol{W}\boldsymbol{A}_2 = \boldsymbol{0}$; $\boldsymbol{Z}\boldsymbol{A}_3 = \boldsymbol{Z}$ and $\boldsymbol{Z}\boldsymbol{A}_4 = \boldsymbol{Z}$. Hence, $\boldsymbol{A}_1 = \boldsymbol{A}_3 = \boldsymbol{A}_4 = \boldsymbol{I}_k$ and $\boldsymbol{A}_2 = \boldsymbol{0}$, so that the matrix $\boldsymbol{A}$ becomes $\begin{bmatrix} \boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{I}_k & \boldsymbol{I}_k \end{bmatrix}$. c) It is not difficult to see that $\boldsymbol{A}$ is indeed invertible. In fact, $\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{I}_{2k}$ gives us $\boldsymbol{A}^{-1} = \begin{bmatrix} \boldsymbol{I}_k & \boldsymbol{0} \\ -\boldsymbol{I}_k & \boldsymbol{I}_k \end{bmatrix}$. d) The simplest way to show that both fitted values and residuals are the same is to show that the projection matrices for Model 1 and Model 2 are the same. Let $\boldsymbol{P}_1$ be the projection matrix for Model 1 and $\boldsymbol{P}_2$ the one for Model 2. Then,

$$
\begin{aligned}
\boldsymbol{P}_2 &= \boldsymbol{X}_2(\boldsymbol{X}_2'\boldsymbol{X}_2)^{-1}\boldsymbol{X}_2' \\
&= (\boldsymbol{X}_1\boldsymbol{A})[(\boldsymbol{X}_1\boldsymbol{A})'(\boldsymbol{X}_1\boldsymbol{A})]^{-1}(\boldsymbol{X}_1\boldsymbol{A})' \\
&= (\boldsymbol{X}_1\boldsymbol{A})[\boldsymbol{A}'\boldsymbol{X}_1'\boldsymbol{X}_1\boldsymbol{A}]^{-1}\boldsymbol{A}'\boldsymbol{X}_1' \\
&= \boldsymbol{X}_1\boldsymbol{A}\boldsymbol{A}^{-1}(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}(\boldsymbol{A}')^{-1}\boldsymbol{A}'\boldsymbol{X}_1' \\
&= \boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1' \\
&= \boldsymbol{P}_1.
\end{aligned}
$$