Modern methods of data analysis
Homework 3

Vitaliy Pozdnyakov

Problem 1. Let $\lambda > 0$ and X be a real-valued random variable with Poisson distribution with parameter λ: $\forall k = 0, 1, 2, \dots$ $\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$. Show that, for all $t > 0$, $\mathbb{P}(X \geq \lambda + t) \leq \exp\left(-\lambda h\left(\frac{t}{\lambda}\right)\right)$, where $h(u) = (1+u)\log(1+u) - u$.

Solution. First, use the Chernoff bound

$$\mathbb{P}(X \geq \lambda + t) \leq \frac{\mathbb{E}\left[e^{mx}\right]}{e^{m(\lambda+t)}} \text{ for all } m > 0.$$

Next, rewrite the expectation as

$$\mathbb{E}\left[e^{mx}\right] = \sum_{k=0}^{\infty} e^{mk} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(e^m \lambda)^k}{k!}.$$

Next, use the property of the exponential function $e^x = \sum_{i=0}^{\infty} \frac{x^k}{k!}$ and then

$$\mathbb{E}\left[e^{mx}\right] = \exp(e^m \lambda - \lambda).$$

Thereby, we obtain

$$\mathbb{P}(X \geq \lambda + t) \leq \frac{\exp(e^m \lambda - \lambda)}{\exp(m\lambda + mt)} = \exp[e^m \lambda - \lambda - m(\lambda + t)].$$

From the task description we have $t > 0$ and $\lambda > 0$, so $\log\left(\frac{\lambda+t}{\lambda}\right) > 0$ too. Finally, make a substitution $m = \log\left(\frac{\lambda+t}{\lambda}\right)$ that satisfy the Chernoff bound

$$\begin{aligned}\mathbb{P}(X \geq \lambda + t) &\leq \exp\left[\left(\frac{\lambda+t}{\lambda}\right)\lambda - \lambda - \log\left(\frac{\lambda+t}{\lambda}\right)(\lambda+t)\right] = \exp\left[t - \log\left(\frac{\lambda+t}{\lambda}\right)(\lambda+t)\right] \\ &= \exp\left[-\lambda\left(\left(1+\frac{t}{\lambda}\right)\log\left(1+\frac{t}{\lambda}\right) - \frac{t}{\lambda}\right)\right] = \exp\left(-\lambda h\left(\frac{t}{\lambda}\right)\right).\end{aligned}$$

Problem 2. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let $Y : \Omega \to \mathbb{R}$ and $X : \Omega \to \mathbb{R}^d$ be two random variables.

Task 1. Suppose that $\mathbb{E}[Y^2] < +\infty$. Show that $\mathbb{E}[r(X)^2] < +\infty$.

Solution. Rewrite the expression $\mathbb{E}(r(X)^2)$ via definition of the expectation and regression function

$$\mathbb{E}[r(X)^2] = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} y \mathrm{d}k_x(y)\right)^2 \mathrm{d}P_X(x).$$

Next, let $f(X, Y) = Y^2 + 0 \cdot X$. Use the Fubini's theorem

$$\mathbb{E}[Y^2] = \mathbb{E}[f(X,Y)] = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} y^2 \mathrm{d}k_x(y) \right) \mathrm{d}P_X(x).$$

Now, compare the dissimilar parts of the integrals. Since the function $g(x) = x^2$ is convex we can apply the Jensen's theorem

$$\left(\int_{\mathbb{R}} y \mathrm{d}k_x(y) \right)^2 \leq \int_{\mathbb{R}} y^2 \mathrm{d}k_x(y).$$

And thereby,

$$\mathbb{E}[r(X)^2] \leq \mathbb{E}[Y^2] < +\infty.$$

Task 2. Suppose that $\mathbb{E}[Y^2] < +\infty$. Show that, for any $f : \mathbb{R}^d \to \mathbb{R}$ such that $\mathbb{E}[f(X)^2] < +\infty$, we have $\mathbb{E}[(Y - f(X))^2] = \mathbb{E}[(Y - r(X))^2] + \mathbb{E}[(r(X) - f(X))^2]$.

Solution. First, use the Fubini's theorem

$$\mathbb{E}[(Y - f(X))^2] = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} (y - f(x))^2 \mathrm{d}k_x(y) \right) \mathrm{d}\mathbb{P}_X(x).$$

Next, add and subtract the regression function $r(x)$

$$\mathbb{E}[(Y - f(X))^2] = \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} (y - r(x) + r(x) - f(x))^2 \mathrm{d}k_x(y) \right) \mathrm{d}\mathbb{P}_X(x).$$

Substitute $a = y - r(x)$ and $b = r(x) - f(x)$ and rewrite the integral with respect to polynomial decomposition $(a - b)^2 = a^2 - 2ab + b^2$

$$\mathbb{E}[(Y - f(X))^2] = \int_{\mathbb{R}^d} \int_{\mathbb{R}} a^2 \mathrm{d}k_x(y) \mathrm{d}\mathbb{P}_X(x) - 2 \int_{\mathbb{R}^d} \int_{\mathbb{R}} ab \cdot \mathrm{d}k_x(y) \mathrm{d}\mathbb{P}_X(x) + \int_{\mathbb{R}^d} \int_{\mathbb{R}} b^2 \mathrm{d}k_x(y) \mathrm{d}\mathbb{P}_X(x).$$

Consider the summands separately. First summand can be rewrited as

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}} (y - r(x))^2 \mathrm{d}k_x(y) \mathrm{d}\mathbb{P}_X(x) = \mathbb{E}[(Y - r(X))^2].$$

Second summand can be reduced because

$$2 \int_{\mathbb{R}^d} \int_{\mathbb{R}} (y - r(x))(r(x) - f(x)) \mathrm{d}k_x(y) \mathrm{d}\mathbb{P}_X(x) = 2 \int_{\mathbb{R}^d} (r(x) - f(x)) \left(\int_{\mathbb{R}} (y - r(x)) \mathrm{d}k_x(y) \right) \mathrm{d}\mathbb{P}_X(x)$$

and the expression in the big brackets is

$$\int_{\mathbb{R}} (y - r(x)) \mathrm{d}k_x(y) = \int_{\mathbb{R}} y \cdot \mathrm{d}k_x(y) - r(x) = 0.$$

Third summand is represented as

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}} (r(x) - f(x))^2 \mathrm{d}k_x(y) \mathrm{d}\mathbb{P}_X(x) = \mathbb{E}[(r(X) - f(X))^2].$$

Finally, union the summands into the result

$$\mathbb{E}[(Y - f(X))^2] = \mathbb{E}[(Y - r(X))^2] + \mathbb{E}[(r(X) - f(X))^2].$$

Task 3. Suppose that $Y = \varphi(X) + Z$, for some function $\varphi : \mathbb{R}^d \to \mathbb{R}$ where Z is independent of X and satisfies $\mathbb{E}[Z] = 0$ and $\mathbb{E}[Z^2] = \sigma^2 \in (0, +\infty)$.

Subtask 3.a. Show that φ is the regression function of Y given X.

Solution. Via another definition of the regression function

$$r(x) := \mathbb{E}[Y|X = x] = \mathbb{E}[\varphi(x) + Z|X = x] = \mathbb{E}[\varphi(x)|X = x] + \mathbb{E}[Z|X = x].$$

Since Z is independent of X then $\mathbb{E}[Z|X = x] = \mathbb{E}[Z] = 0$. Next, we can consider φ as a constant C_i for each given x_i and thereby $\mathbb{E}[\varphi(x)|X = x] = \varphi(x)$. Hence, $r(x) = \varphi(x)$.

Subtask 3.b. Show that, for any $f : \mathbb{R}^d \to \mathbb{R}$ such that $\mathbb{E}[f(X)^2] < +\infty$, we have $\mathbb{E}[(Y - f(X))^2] \geq \sigma^2$.

Solution. In the Task 2 we obtained that

$$\mathbb{E}[(Y - f(X))^2] = \mathbb{E}[(Y - r(X))^2] + \mathbb{E}[(r(X) - f(X))^2].$$

In addition, from the previous subtask we know that $r(x) = \varphi(x)$, so

$$\mathbb{E}[(Y - f(X))^2] = \mathbb{E}[(Y - \varphi(X))^2] + \mathbb{E}[(\varphi(X) - f(X))^2].$$

Next, substitute $\varphi(X) = Y - Z$ in the first brackets in the right side

$$\mathbb{E}[(Y - f(X))^2] = \mathbb{E}[Z^2] + \mathbb{E}[(\varphi(X) - f(X))^2].$$

Obviously $(\varphi(X) - f(X))^2 \geq 0$ and then $\mathbb{E}[(\varphi(X) - f(X))^2] \geq 0$ too. Thereby

$$\mathbb{E}[(Y - f(X))^2] \geq \mathbb{E}[Z^2] = \sigma^2.$$