Stat 513
Fall 2025
Problem Set 4
Topic 4: Conditional Distributions and Conditional Expectation

Due Sunday, October 19 at 23:59

Problem 1. Tightness of Chebyshev's Inequality

Consider a random variable X with range $\{-a, 0, a\}$ for some $a > 0$ such that:

$$P(X = -a) = p$$

$$P(X = 0) = 1 - 2p$$

$$P(X = a) = p.$$

Show that for some value of k, Chebyshev's inequality holds with equality.

Problem 1 Solution

Recall that Chebyshev's inequality is:

$$P\left(|X - \mathbb{E}[X]| \geq k\sigma\right) \leq \frac{1}{k^2}$$

We can see that

$$\mathbb{E}[X] = (-a)p + 0 + ap = 0$$

and

$$\sigma^2 = \mathbb{E}[X^2] = (-a)^2 p + 0 + a^2 p = 2a^2 p$$

so $\sigma = \sqrt{2p}a$.

Let $k = \frac{1}{\sqrt{2p}}$. Then Chebyshev's inequality becomes

$$P\left(|X| \geq a\right) \leq 2p.$$

Evaluating the probability on the left-hand side, we get

$$P\left(|X| \geq a\right) = P(X = a) + P(X = -a) = p + p = 2p$$

Problem 2. Cantelli's inequality

i) Prove the following inequality:

$$P(X - \mathbb{E}[X] \geq \lambda) \leq \frac{\sigma^2}{\sigma^2 + \lambda^2}$$

where $\lambda \geq 0$.

ii) When is Cantelli's inequality better than Chebyshev's inequality?

Problem 2 Solution

i) Let $Y = X - \mathbb{E}[X]$, then for any $a > 0$,

$$P(Y \geq \lambda) = P(Y + a \geq \lambda + a) \leq P((Y + a)^2 \geq (\lambda + a)^2).$$

Since $(Y + a)^2$ is non-negative, by Markov's inequality,

$$P((Y + a)^2 \geq (\lambda + a)^2) \leq \frac{\mathbb{E}\left[(Y + a)^2\right]}{(\lambda + a)^2} = \frac{\sigma^2 + a^2}{(\lambda + a)^2}.$$

Since the left-hand side is less than or equal to the right-hand side for all $a > 0$, we can obtain a tighter bound by minimizing the right-hand side over a.

We can then obtain Cantelli's inequality by minimizing this bound over a:

$$P(Y \geq \lambda) \leq \frac{\sigma^2}{\sigma^2 + \lambda^2}$$

ii) Cantelli's inequality is better than Chebyshev's inequality when evaluating $P(X - \mathbb{E}[X] \geq \lambda)$, since

$$P\left(X - \mathbb{E}[X] \geq \lambda\right) \leq P\left(|X - \mathbb{E}[X]| \geq \lambda\right) \leq \frac{\sigma^2}{\lambda^2}$$

which is strictly greater than $\frac{\sigma^2}{\lambda^2 + \sigma^2}$ since $\sigma^2 > 0$. When evaluating $P\left(|X - \mathbb{E}[X]|\right)$, Chebyshev's inequality is better since applying Cantelli's inequality to both sides yields

$$P\left(|X - \mathbb{E}[X]| \geq \lambda\right) \leq \frac{2\sigma^2}{\sigma^2 + \lambda^2}$$

which is strictly greater than $\frac{\sigma^2}{\lambda^2}$ when $\lambda > \sigma$.

Problem 3. A Sum Rule for Expectations

Show that if the range of X is the natural numbers, then

$$\mathbb{E}[X] = \sum_{n=1}^{\infty} P(X \geq n).$$

Problem 3 Solution

Since X is a discrete random variable, we can use the following expression for the expectation:

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} iP(X = i)$$

This becomes

$$\begin{aligned}\mathbb{E}[X] &= \sum_{i=1}^{\infty}\sum_{n=1}^{i} P(X = i) = \sum_{1\le n\le i<\infty} P(X = i)\\ &= \sum_{n=1}^{\infty}\sum_{i=n}^{\infty} P(X = i) = \sum_{n=1}^{\infty} P(X \ge n)\end{aligned}$$

since $\sum_{i=n}^{\infty} P(X = i) = P(X \ge n)$.

Problem 4. Conditional Densities for Absolutely Continuous Distributions

Let X and Y denote real-valued random variables such that the distribution of (X, Y) is absolutely continuous with density function

$$p(x, y) = \frac{1}{x^3y^2}, \quad x > 1,\ y > 1/x.$$

Find conditional distributions for X given $Y = y$ and Y given $X = x$.

Problem 4 Solution The marginal density function of X is

$$p_X(x) = \int_{1/x}^{\infty} \frac{1}{x^3y^2} dy = -\frac{1}{x^3y}\bigg|_{1/x}^{\infty} = \frac{1}{x^2}.$$

The marginal density function of Y is

$$p_Y(y) = \int_{1}^{\infty} \frac{1}{x^3y^2} dy = -\frac{1}{2x^2y^2}\bigg|_{1}^{\infty} = \frac{1}{2y^2}.$$

Since these are absolutely continuous distributions, we can use Sevirini Theorem 2.3 to get

$$p_{X|Y}(x|y) = \frac{p(x, y)}{p_Y(y)} = \frac{2}{x^3} \quad \text{when } x > 1, y > 1/x$$

and

$$p_{Y|X}(y|x) = \frac{p(x, y)}{p_X(y)} = \frac{1}{xy^2} \quad \text{when } x > 1, y > 1/x$$

Problem 5. Mixed Distribution Let (X, Y) denote a two-dimensional random vector with range $(0, \infty) \times \{1, 2\}$ such that for any set $A \subset (0, \infty)$ and $y \in \{1, 2\}$,

$$P(X \in A, Y = y) = \frac{1}{2}\int_A y\exp(-yx)dx.$$

Find conditional distributions for X given $Y = y$ and for Y given $X = x$.

Problem 5 Solution Consider the distribution of X given $Y = y$. The marginal distribution of Y is discrete, with

$$P(Y = 1) = P(X \in \mathcal{X}, Y = 1) = \frac{1}{2}\int_0^\infty y\exp(-yx)\,dx = \frac{1}{2}$$

so it must be that $P(Y = 2) = \frac{1}{2}$ as well. Since $P(Y = y)$ is non-zero for $y \in \{1, 2\}$, we can evaluate the conditional distribution function directly:

$$P(X \in A|Y = y) = \frac{P(X \in A, Y = y)}{P(Y = y)} = \int_A y\exp(-yx)\,dx.$$

It follows that conditional on $Y = y$, X has an absolutely continuous distribution with density function $y\exp(-yx)$.

For the distribution of Y given X, we are looking for a function $q(B, x)$ that satisfies the relationship

$$P(X \in A, Y \in B) = \int_A q(B, x)\,dF_X(x) = \frac{1}{2}\int_A y\exp(-xy)\,dx$$

where $B \subset \{1, 2\}$. We can see that the marginal distribution for X is

$$P(X \in A) = P(X \in A, Y \in \{1, 2\}) = P(X \in A, Y = 1) + P(X \in A, Y = 2) = \int_A \frac{1}{2}\left(\exp(-x) + 2\exp(-2x)\right)dx.$$

It follows that

$$P(X \in A, Y \in B) = \int_A q(B, x)\,dF_X(x) = \frac{1}{2}\int_A q(B, x)\left(\exp(-x) + 2\exp(-2x))\right)dx = \frac{1}{2}\int_A y\exp(-yx)\,dx$$

In order to satisfy the last equality in the above, we can set $q(B, x)$ (for $B = \{y\}$) to be

$$P(Y = y|X = x) = \frac{y\exp(-yx)}{\exp(-x) + 2\exp(-2x)}.$$

Problem 6. Non-uniqueness of Conditional Probabilities Using the joint distribution from the previous problem, describe two conditional distributions (i.e., set functions $q_1(\cdot, y)$ and $q_2(\cdot, y)$ that satisfy the definition of conditional probability) that differ on an uncountable set.

Problem 6 Solution Any example of conditional densities that differ on an uncountable set of probability zero (there are many such sets).

Problem 7. Conditional Distributions as a Limit

Let X and Y denote real-valued random variables such that the distribution of (X, Y) is absolutely continuous with density function f, and let f_X denote the marginal density function of X. Suppose that there exists a point x_0 such that $f_X(x_0) > 0$, f_X is continuous at x_0, and for almost all y, $f(\cdot, y)$ is continuous at x_0. Let $A \subset \mathbb{R}$. For each $\epsilon > 0$, let

$$d(\epsilon) = P(Y \in A | x_0 \leq X \leq x_0 + \epsilon).$$

Show that

$$P(Y \in A | X = x_0) = \lim_{\epsilon \to 0} d(\epsilon).$$

Problem 7 Solution We can see that

$$d(\epsilon) = \frac{P(X \in [x_0, x_0 + \epsilon], Y \in A)}{P(X \in [x_0, x_0 + \epsilon])} = \frac{\int_{x_0}^{x_0+\epsilon} \int_A f_{X,Y}(x, y)\, dy\, dx}{\int_{x_0}^{x_0+\epsilon} \int_{\mathcal{Y}} f_{X,Y}(x, y)\, dy\, dx}.$$

Taking the limit as $\epsilon \to 0$, we get

$$\lim_{\epsilon \to 0} d(\epsilon) = \frac{\int_A f_{X,Y}(x_0, y)\, dx}{\int_{\mathcal{Y}} f_{X,Y}(x_0, y)\, dy} = \int_A \frac{f_{X,Y}(x_0, y)}{f_X(x_0)}\, dx.$$

given that $f_X(x_0) > 0$. Since the distributions are absolutely continuous, we can use Sevirini Theorem 2.3 to infer that $f_{Y|X=x_0}(y|x_0) = \frac{f_{X,Y}(x_0,y)}{f_X(x_0)}$ is the density function for the distribution of $X|Y$, which means that $P(Y \in A|X = x_0) = \lim_{\epsilon \to 0} d(\epsilon)$ as required.

Problem 8. Sums of Conditional Expectations

Let X denote a real valued random variable with range $\mathcal{X}$, such that $E[|X|] < \infty$. Let $A_1, \ldots, A_n$ denote disjoint subsets of $\mathcal{X}$. Show that

$$E(X) = \sum_{i=1}^{N} \mathbb{E}[X|X \in A_j] P(X \in A_j).$$

Problem 8 Solution Let $Y_j = 1\{X \in A_j\}$; note that this is a random variable that takes values zero and one, and that $\{Y_j = 1\} = \{X \in A_j\}$. By definition, we know that the conditional expectation $E[X|Y_j]$ must satisfy

$$\mathbb{E}[X 1\{Y \in B\}] = \int_B \mathbb{E}[X|Y_j = y] dF_Y(y).$$

Setting $B = \{1\}$, we can see that

$$\mathbb{E}[X1\{X \in A_j\}] = \int_{y=1} \mathbb{E}[X|Y_j = 1] dF_Y(y) \tag{1}$$

$$= \mathbb{E}[X|Y_j] P(Y_j = 1) \tag{2}$$

$$= \mathbb{E}[X|A_j] P(X \in A_j). \tag{3}$$

Thus,

$$\sum_{i=1}^{n} \mathbb{E}[X|A_j] P(X \in A_j) = \sum_{i=1}^{n} \mathbb{E}[X1\{X \in A_j\}] = \mathbb{E}\left[X \sum_{j=1}^{n} 1\{X \in A_j\}\right] = E[X]$$

Since $\sum_{j=1}^{n} 1\{X \in A_j\} = 1$.