Wenxiao Yang

# Homework 7

## Homework Policy

- You must turn in your own work, do not copy answers from others. This is not helpful to you at all.
- You are welcome and encouraged to work together on homework. If you work with others, write the following on the top of your assignment: I worked with .... on this assignment.
- I am fully aware that it is possible to use online resources and merely copy the derived solutions. This does not help you to learn the material and is going to be detrimental for you in the exams.
- Show all supporting work.

## Concepts:

- Conditional Expectation and Variance
- Cauchy Schwarz Inequality, Jensen's Inequality
- Markov, Chebyshev and Chernoff Inequalities.

## Exercises

1. **Conditional Expectation and Variance**(3 points)

An insurance company covers disasters in two neighboring regions, $R_1$ and $R_2$. Let $I_1$ and $I_2$ be the indicator random variables for whether $R_1$ and $R_2$ are hit by the insured disaster, respectively. The indicators $I_1$ and $I_2$ may be dependent. Let $p_j = \mathbb{E}(I_j)$ for $j = 1, 2$ and $p_{12} = E(I_1 I_2)$. The company reimburses a total cost of

$$C = I_1 T_1 + I_2 T_2$$

to these regions, where $T_j$ has mean $\mu_j$ and variance $\sigma_j^2$. Assume that $T_1$ and $T_2$ are independent of each other and independent of $I_1, I_2$.

(a) Find $E(C)$. (1 point)

(b) Find $Var(C)$. (2 points)

(Hint: You can define $\mathbb{E}(C|I_1, I_2)$ which is a random variable and is a function $g(I_1, I_2)$ where for each possible value of $I_1 = i_1, I_2 = i_2$ the function $g(i_1, i_2) = \mathbb{E}(C|I_1 = i_1, I_2 = i_2)$. Now use law of total expectation and total variance.

2. **AM GM inequality using Jensen's inequality**(2 points)

   The famous arithmetic mean geometric mean inequality says that for any positive numbers $a_1, \ldots, a_n$,
   $$\frac{a_1 + \cdots + a_n}{n} \geq (a_1 a_2 \ldots a_n)^{1/n}.$$
   Show that this inequality follows from Jensen's inequality by considering $\mathbb{E} \log X$ for a random variable $X$ whose possible values are $a_1, \ldots, a_n$.

3. **Fill in the blank inequalities**(3 points, 0.5 points per part)

   Let $X$ and $Y$ be i.i.d continuous random variables. Assume that the various expected values exist. Write the most appropriate of $\leq, \geq, =$ or ? in the blank for each part. Give reasoning for every part.

   (a) $\exp(-\mathbb{E}(X))$ ____ $\mathbb{E}(\exp(-X))$

   (b) $P(X > Y + 3)$ ____ $P(Y > X + 3)$

   (c) $P(X > Y + 3)$ ____ $P(X > Y - 3)$

   (d) $\mathbb{E}(X^4)$ ____ $\big(\mathbb{E}(XY)\big)^2$

   (e) $Var(Y)$ ____ $\mathbb{E}Var(Y|X)$

   (f) $P(|X + Y| > 3)$ ____ $\mathbb{E}|X|$

4. **Birthday Problem**(3 points)

   Suppose there are 10 people in a party. Give a lower bound on the probability that there exists one pair of people with birthdays at most two days apart.

   (Hint: Use the second moment inequality $P(X = 0) \leq \frac{VarX}{\mathbb{E}X^2}$ where $X$ is a non negative random variable. Also introduce indicator random variables for each pair of people indicating whether their birthdays are atmost two days apart or not. A similar exercise was done in the lecture.)

   *The maximum you can score in this homework is 11. This homework is worth 10 points in the overall scheme of things as will be every homework. If you score 11 you will earn 1 extra credit point.*

# 1 Survey

Now that we are halfway through the course and the first midterm is over, I thought I will be asking a few questions here regarding the course. Feel free to answer these along with this homework so that I can get some feedback. Also feel free to reach out to me by email if you have anything specific to talk about.

1. How was the midterm exam? The exams in the future will be of a similar level. If you go through the homeworks, additional problems and the practice midterm you will

stand a good chance of doing well in the exams. The practice exam I release before the exams will bear similarity with the actual exam.

2. Do the Office Hour Timings work for you? I see very few people coming for the Tuesday office hours from 8 pm to 9 pm. I have tried to accommodate most time zones in the world. If there is a majority of you wanting different OH times I can change the timing.

3. Are the lectures understandable? So far, I am following the treatment in Blitzstein's book and it will help you if you read the assigned sections. I know the material could be hard for some of you and if so I would encourage you to write your doubts on Compass or reach out to me or the TA during OH.

4. Feel free to write to me or the TA on any other matter that could improve your experience of this course.

1. (a) $E(c) = E(E(c \mid T_1, T_2))$

$$= E(p_1 T_1 + p_2 T_2)$$

$$= p_1 \mu_1 + p_2 \mu_2$$

(b) $Var(c) = E(Var(C \mid I_1, I_2)) + Var(E(C \mid I_1, I_2))$

$$= E(I_1^2 \sigma_1^2 + I_2^2 \sigma_2^2) + Var(\mu_1 I_1 + \mu_2 I_2)$$

$$= p_1 \sigma_1^2 + p_2 \sigma_2^2 + \mu_1 (p_1 - p_1^2) + \mu_2 (p_2 - p_2^2)$$

2. Since $g(x) = \log X$ is concave, we know

$$E \log(X) \le \log(E(X))$$

assume r.v. $X$ possible values are $a_1, \dots, a_n$, and $P(X = a_i) = \frac{1}{n}$
$i = 1, 2, \dots n.$

Hence $E \log X = \frac{1}{n}(\log a_1 + \log a_2 + \cdots + \log a_n)$

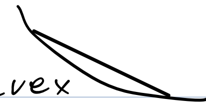
$$= \log(a_1 a_2 \cdots a_n)^{\frac{1}{n}}$$

$$\log E(X) = \log\left(\frac{1}{n}(a_1 + a_2 + \cdots + a_n)\right)$$

Because $E \log(X) \le \log(E(X))$

$$\log(a_1 a_2 \cdots a_n)^{\frac{1}{n}} \le \log\left(\frac{1}{n}(a_1 + a_2 + \cdots + a_n)\right)$$

$\log(x)$ is monotone increasing $\Rightarrow (a_1 a_2 \cdots a_n)^{\frac{1}{n}} \le \frac{1}{n}(a_1 + a_2 + \cdots + a_n).$

3. (a) $\leq$ $\qquad g(x) = e^{-x}$ is convex 

(b) $=$ $\qquad P(X-Y>3) = P(Y-X>3)$

(c) $\leq$ $\qquad P(X-Y>3) \leq P(X-Y>-3)$

(d) $\geq$ $\qquad E^2(XY) \leq E(X^2)E(Y^2) = E^2(X^2) \leq E(X^4)$

(e) $\geq$ $\qquad E\,Var(Y|X) + \underbrace{Var(E(Y|X))}_{\geq 0} = Var(Y)$

(f) $\leq$ $\qquad P(|X+Y|>3) \leq \dfrac{E|X+Y|}{3} \leq \dfrac{E|X|+E|Y|}{3} = \underset{\leq E|X|}{\dfrac{2E|X|}{3}}$

4. $X_{ij}$ is the r.v. such that $X_{ij} = \begin{cases} 1, & i \text{ and } j\text{'s birthdays are at most two days apart.} \\ 0, & \text{more than two days.} \end{cases}$

$$P(X_{ij}=0) \leq \frac{Var X_{ij}}{E X_{ij}^2} = \frac{E X_{ij}^2 - E^2 X_{ij}}{E X_{ij}^2} = 1 - \frac{E^2 X_{ij}}{E X_{ij}^2} = 1 - E X_{ij}$$

$$= 1 - \left(1 \times \frac{5}{365} + 0 \cdot \frac{360}{365}\right) = \frac{72}{73}$$

$$1 - \prod_{j=2}^{10} \prod_{i=1}^{j-1} P(X_{ij}=0) \geq 1 - \left(\frac{72}{73}\right)^{45}$$

lower bound