

Math, Problem Set #1, Probability Theory

OSM Lab, Karl Schmedders

Due Monday, June 26 at 8:00am

1. **Exercises from chapter.** Do the following exercises in Chapter 3 of ? : 3.6, 3.8, 3.11, 3.12 (watch this movie [clip](#)), 3.16, 3.33, 3.36.

Solutions:

3.6

Assumptions:

(a) $\Omega = \cup_{i \in I} B_i$

(b) $B_i \cap B_j = \emptyset$

Proof:

$$P(A) = P(\Omega \cap A) = P(\cup_{i \in I} B_i \cap A) = P(\cup_{i \in I} (B_i \cap A)) = \sum_{i \in I} P(A \cap B_i)$$

3.8

First, if the events E_k are independent, then their complements E_k^c are independent, too.

$$\begin{aligned} 1 - \prod_{k=1}^n (1 - P(E_k)) &= 1 - \prod_{k=1}^n (P(E_k)^c) \\ &= 1 - P(\cap_{k=1}^n E_k^c) \end{aligned}$$

and by De Morgan's law we get

$$\begin{aligned} &= 1 - P((\cup_{k=1}^n E_k)^c) \\ &= P(\cap_{k=1}^n E_k) \end{aligned}$$

3.11

We are interested in $P(s = \text{crime} | \text{test} +)$, given

- $P(s = \text{crime}) = \frac{1}{250 \text{ million}}$ because one person out of the total population committed the crime
- $P(\text{test} +) = \frac{1}{3 \text{ million}}$
- I assume the test is 100% reliable in the sense that $P(\text{test} + | s = \text{crime}) = 1$.

- In order to allow for the possibility of testing positive the s has to be in the random sample of 1 million DNA samples so $P(s \text{ in the sample} | s = \text{crime}) = \frac{1}{250}$

Then the conditional probability

$$\begin{aligned}
 P(s = \text{crime} | \text{test} +) &= \frac{P(\text{test} + | s = \text{crime})P(s = \text{crime})}{P(\text{test} +)} \\
 &= \frac{P(s \text{ in the sample} | s = \text{crime})P(s = \text{crime})}{P(\text{test} +)} \\
 &= \frac{\frac{1}{250} \times \frac{1}{250 \text{ million}}}{\frac{1}{3 \text{ million}}} \\
 &= \frac{3}{250^2}
 \end{aligned}$$

3.12

- i not-shifting strategy. When one decides before the game starts on not to shift after Hall reveals a goat the probability of winning is by randomly picking the winning door is simply $P(W) = \frac{1}{3}$.
- ii shifting strategy. Hall will always reveal a goat. So if one has picked the winning door in the first round one will loose with probability $\frac{1}{3} \times 1$ because one will pick a goat by switching in the second round. In contrast, if one has chosen a goat door in the first round and Hall reveals the second goat, there is only the winning door left so one wins with probability $\frac{2}{3}$ that is equal to the probability of choosing a goat in the first round.
- iii When there are 10 doors and Hall reveals 8 goats after one has picked a door in the first round it works the same. The probabilities of winning and loosing can be flipped around by following the switching strategy compared to bot switching. $P(W) = \frac{1}{10}$ without switching. With the switching strategy one chooses a goat with probability $\frac{9}{10}$, then Hall reveals the remaining 8 goats and one switches to the winning door. So the probability of winning is $\frac{9}{10}$ in this case.

3.16

$$\begin{aligned}
 \text{Var}[X] &= E[(X - \mu)^2] = E[X^2 - 2X\mu + \mu^2] \\
 &= E[X^2] - 2E[X]\mu + \mu^2 = E[X^2] - 2\mu^2 + \mu^2 \\
 &= E[X^2] - \mu^2
 \end{aligned}$$

3.33

Since $B \sim \text{Binom}(n, p)$, $E[B] = np$ and $\text{Var}[B] = \sigma^2 = p(1 - p)$

$$P\left(\left|\frac{B}{n} - p\right| \geq \varepsilon\right) = P(|B - np| \geq n\varepsilon)$$

And by the Chebyshev's Inequality we get

$$\begin{aligned} P(|B - np| \geq n\varepsilon) &\leq \frac{np(1-p)}{n^2\varepsilon^2} \\ &\leq \frac{p(1-p)}{n\varepsilon^2} \end{aligned}$$

3.36

Since $X_i \sim \text{Bernoulli}(p)$ for $i \in (1, 6242)$ with $E[x] = p = 0.801$ and $\text{Var}[X] = \sigma^2 = p(1 - p) = 0.199 \times 0.801$, let the number of students actually enrolling be $S = \sum 6242_i = 1X_i$, then the variable $\frac{S-6242p}{\sigma\sqrt{6242}} \sim N(0, 1)$. Then the probability that more than 5500 students will enroll is given by $1 - P(x < \frac{5500-6242p}{\sigma\sqrt{6242}}) = 0$. This probability is practically zero.

2. Construct examples of events A , B , and C , each of probability strictly between 0 and 1, such that

- (a) $P(A \cap B) = P(A)P(B)$, $P(A \cap C) = P(A)P(C)$, $P(B \cap C) = P(B)P(C)$, but $P(A \cap B \cap C) \neq P(A)P(B)P(C)$.
- (b) $P(A \cap B) = P(A)P(B)$, $P(A \cap C) = P(A)P(C)$, $P(A \cap B \cap C) = P(A)P(B)P(C)$, but $P(B \cap C) \neq P(B)P(C)$. (Hint: You can let Ω be a set of eight equally likely points.)

Solutions:

- (a) let $\Omega = 1, 2, 3, 4, 5, 6, 7, 8$ with $P(x = x_i) = \frac{1}{8} \forall x_i \in \Omega$ and with the events

$$\begin{aligned} A &= \{1, 2, 3, 4\} \\ B &= \{1, 2, 5, 6\} \\ C &= \{3, 4, 5, 6\} \end{aligned}$$

Such that $P(A) = P(B) = P(C) = \frac{4}{8}$ and the intersections $A \cap B$, $A \cap C$, $B \cap C$ each have two elements leading to $P(A \cap B) = P(A \cap C) = P(B \cap C) = P(A)P(C) = P(A)P(B) = P(B)P(C) = \frac{2}{8}$. Furthermore, $P(A \cap B \cap C) = P(\emptyset) = 0 \neq P(A)P(B)P(C) = \frac{1}{8}$. \emptyset

(b) for the same Ω as in (a) be

$$A = \{1, 2, 3, 4\}$$

$$B = \{3, 4, 6, 7\}$$

$$C = \{1, 2, 4, 5\}$$

Such that $P(A) = P(B) = P(C) = \frac{4}{8}$ and the intersections $A \cap B$ and $A \cap C$ each have two elements leading to $P(A \cap B) = P(A \cap C) = P(A)P(C) = P(A)P(B) = \frac{2}{8}$. Furthermore, $P(B \cap C) = P(x = 4) = \frac{1}{8} \neq P(B)P(C) = \frac{2}{8}$ and $P(A \cap B \cap C) = P(4) = \frac{1}{8} = P(A)P(B)P(C)$.

3. Prove that Benford's Law is, in fact, a well-defined discrete probability distribution.

Solutions: Benford's law holds for the sample space

1. since $2 \geq (1 + \frac{1}{d}) \geq (1 + \frac{1}{9}) \forall d \in \Omega$ it follows that $0 \geq \log_{10}(1 + \frac{1}{d}) \geq 1$ holds.

2.

$$\begin{aligned} \sum_{d=1}^9 \log_{10} \left(1 + \frac{1}{d} \right) &= \log_{10} \left(\prod_{d=1}^9 1 + \frac{1}{d} \right) \\ &= \log_{10}(2 \times 1.5 \times 1.33 \times 1.25 \times 1.2 \times 1.167 \times 1.143 \times 1.125 \times 1.111 \times) \\ &= \log_{10}(10) = 1 \end{aligned}$$

3. for any events d_i and d_j with $i \neq j$, the intersection $d_i \cap d_j = \emptyset$ (i.e. they are pairwise disjoint) such that $P(d_i \cup d_j) = P(d_i) + P(d_j)$
4. A person tosses a fair coin until a tail appears for the first time. If the tail appears on the n th flip, the person wins 2^n dollars. Let the random variable X denote the player's winnings.

(a) (St. Petersburg paradox) Show that $E[X] = +\infty$.

(b) Suppose the agent has log utility. Calculate $E[\ln X]$.

Let X be the payoff from the St. Petersburg game, then its expected value can be calculated by

$$E[X] = \sum_{k=1}^{\infty} \left(\frac{1}{2} \right)^k 2^k = 1 + 1 + 1 + 1 + \dots = +\infty$$

When the payoff of one game is weighted with the utility function $u(x) = \ln(x)$, the expected utility is given by

$$E[u(X)] = \sum_{k=1}^{\infty} \left(\frac{1}{2} \right)^k \ln(2^k) = \ln(2) \sum_{k=1}^{\infty} \frac{k}{2^k} = \ln(2)S$$

where

$$\begin{aligned} S &= \sum_{k=1}^{\infty} \frac{k}{2^k} = S - \frac{1}{2}S + \frac{1}{2}S = S - \sum_{k=1}^{\infty} \frac{1}{2} \frac{k}{2^k} + \frac{1}{2}S = S - \sum_{k=1}^{\infty} \frac{k}{2^{k+1}} + \frac{1}{2}S \\ &= \sum_{k=1}^{\infty} \frac{k}{2^k} - \sum_{k=2}^{\infty} \frac{k-1}{2^k} + \frac{1}{2}S = \sum_{k=1}^{\infty} \frac{1}{2^k} + \frac{1}{2}S = 1 + \frac{1}{2}S \Rightarrow S = 2 \end{aligned}$$

so,

$$E[X] = 2\ln(2)$$

5. (Siegel's paradox) Suppose the exchange rate between USD and CHF is 1:1. Both a U.S. investor and a Swiss investor believe that a year from now the exchange rate will be either 1.25 : 1 or 1 : 1.25, with each scenario having a probability of 0.5. Both investors want to maximize their wealth in their respective home currency (a year from now) by investing in a risk-free asset; the risk-free interest rates in the U.S. and in Switzerland are the same. Where should the two investors invest?

Solution: Assume an investor invests a unit either in the risk free asset in his home country or abroad. Assume, without loss of generality, that the risk free rate is zero, then $E^{home} = 1$. Whereas

$$E^{abroad} = 0.5(1.25 + \frac{1}{1.25}) = 0.5 \times 2.05 = 1.025 > 1$$

Depending on the risk aversion on the investor, the investor should invest should be willing to take the currency expected return is higher than for the home investment without currency risk.

6. Consider a probability measure space with $\Omega = [0, 1]$.
- Construct a random variable X such that $E[X] < \infty$ but $E[X^2] = \infty$.
 - Construct random variables X and Y such that $P(X > Y) > \frac{1}{2}$ but $E[X] < E[Y]$.
 - Construct random variables X , Y , and Z such that $P(X > Y)P(Y > Z)P(X > Z) > 0$ and $E(X) = E(Y) = E(Z) = 0$.

Solution:

- (a) Let X be a Pareto-distributed random variable with $F(X) = 1 - \left(\frac{x_{min}}{x}\right)^k$ and $1 < k < 2$. The first moment is given by $E[X] = x_{min} \frac{k}{k-1} < +\infty \quad \forall \quad k > 1$ and the second moment is $E[X^2] = x_{min}^2 \frac{k}{k-2} = +\infty \quad \forall \quad k < 2$

- (b) Let X and Y be two random variables that follow the Bernoulli distribution. In particular $P(Y = 1) = \frac{1}{2}$ and $P(Y = 3) = \frac{1}{2}$ so $E[Y] = \frac{1}{2}(1+3) = 1$. Assume that X can take either the value 4 or -4. It holds that $P(X = 4) = P(X > Y) = p$ and be found by

$$\begin{aligned} E[X] &< E[Y](1-p)(-4) + 4p < 1 \\ -4 + 8p &< 1 \\ 8p &< 5 \\ p &< \frac{5}{8} \\ (\text{by assumption}) &\Rightarrow \frac{1}{2} < p < \frac{5}{8} \end{aligned}$$

- (c) Let X , Y and Z be independent and uniformly distributed on $[-1, 1]$ such that $E(X) = E(Y) = E(Z) = 0$. The pairwise joint distributions are also uniform distributions. It follows that $P(X < Y) = P(Y > Z) = P(X > Z) = \frac{1}{2}$ and, hence, $P(X > Y)P(Y > Z)P(X > Z) = \frac{1}{2^3}$.
7. Let the random variables X and Z be independent with $X \sim N(0, 1)$ and $P(Z = 1) = P(Z = -1) = \frac{1}{2}$. Define $Y = XZ$ as the product of X and Z . Prove or disprove each of the following statements.
- (a) $Y \sim N(0, 1)$.
 - (b) $P(|X| = |Y|) = 1$.
 - (c) X and Y are not independent.
 - (d) $Cov[X, Y] = 0$.
 - (e) If X and Y are normally distributed random variables with $Cov[X, Y] = 0$, then X and Y must be dependent.

Solution:

- (a) Proof:

$$\begin{aligned} P(y \leq Y) &= P(y \leq XZ) = P(Z = 1)P(y \leq XZ|Z = 1) + P(Z = -1)P(y \leq XZ|Z = -1) \\ &= 0.5P(y \leq X|Z = 1) + 0.5P(y \leq -X|Z = -1) \\ &= 0.5P(y \leq X) + 0.5P(-y \geq X) \\ &= P(y \leq X) \quad (\text{by symmetry of } X) \end{aligned}$$

and since $X \sim N(0, 1)$ also $Y \sim (0, 1)$.

- (b) i) $Z = 1$ and so it must hold that $X = Y$ so $P(|X| = |Y|)$
 ii) $Z = -1$ and so it must hold that $-X = Y$ so $P(|-X| = |Y|) = P(|X| = |Y|) = 1$

- (c) Independence of X and Y implies that $P(X = x|Y = y) = P(X = x)$. Consider the following counter example: $P(X = 1|Y = 1) = p \neq P(X = 1)P(Y = 1) = p$. Intuitively, they cannot be independent because X is used in construction of Y .

(d)

$$\begin{aligned} Cov[XY] &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY] \\ &= E[X]E[Y] = 0 \end{aligned}$$

Uncorrelated random variables have no *linear* relation. However, this does not generally imply that they must be independent (see(d)).

- (e) False. The other parts of this exercise provide a counter example.

8. Let the random variables X_i , $i = 1, 2, \dots, n$, be i.i.d. having the uniform distribution on $[0, 1]$, denoted $X_i \sim U[0, 1]$. Consider the random variables $m = \min\{X_1, X_2, \dots, X_n\}$ and $M = \max\{X_1, X_2, \dots, X_n\}$. For both random variables m and M , derive their respective cumulative distribution (cdf), probability density function (pdf), and expected value.

Solution:

CDF:

$$\begin{aligned} P(m \leq x) &= P(\text{at least one } X_k \leq x) \\ &= 1 - P(\text{all } X_k > x) \\ &= 1 - (1 - P(\text{all } X_k \leq x)) \\ &= 1 - (1 - P(X_k \leq x))^n \\ &= 1 - (1 - x)^n \quad \text{because } F(X_k) = x \end{aligned}$$

$$P(M \leq x) = P(\text{all } X_k \leq x) = (F(x))^n = x^n$$

PDF:

$$\begin{aligned} f_m(x) &= \frac{dP(m \leq x)}{dx} = n(1 - x)^{n-1} \\ f_M(x) &= \frac{dP(M \leq x)}{dx} = nx^{n-1} \end{aligned}$$

Expectation

$$\begin{aligned} E(m) &= \int_0^1 x f_m(x) dx = \frac{1}{1+n} \\ E(M) &= \int_0^1 x f_M(x) dx = \frac{n}{1+n} \end{aligned}$$

9. You want to simulate a dynamic economy (e.g., an OLG model) with two possible states in each period, a “good” state and a “bad” state. In each period, the probability of both shocks is $\frac{1}{2}$. Across periods the shocks are independent. Answer the following questions using the Central Limit Theorem and the Chebyshev Inequality.

- (a) What is the probability that the number of good states over 1000 periods differs from 500 by at most 2%?
- (b) Over how many periods do you need to simulate the economy to have a probability of at least 0.99 that the proportion of good states differs from $\frac{1}{2}$ by less than 1%?

Solution:

(a)

$$X_k = \text{state}, X \sim \text{Bernoulli}(0.5)$$

$$\mu_X = 0.5, \sigma_X = 0.5$$

$$S = \sum_{k=1}^{1000} X_k, S \sim N(500, 250)$$

We are looking for $P(490 \leq S) + P(S \geq 510) = P(490 \leq S) + (1 - P(S \leq 510)) = 2P(S \leq 490) \approx 0,968$

(b) using the weak law of large numbers we have

$$\begin{aligned} P\left(\left|\frac{\sum^n X_k}{n} - \mu\right| \geq \varepsilon\right) &\leq \frac{\sigma^2}{n\varepsilon^2} \\ \Leftrightarrow 0.99 = 1 - P\left(\left|\frac{\sum^n X_k}{n} - \mu\right| \geq \varepsilon\right) &\geq 1 - \frac{\sigma^2}{n\varepsilon^2} \\ \rightarrow n = \frac{\sigma^2}{0.01\varepsilon^2} \end{aligned}$$

and with $\varepsilon = 0.5 \times 0.01 = 0.005$ we get

$$n = \frac{100}{4 \times 0.005^2} = \frac{100}{0.0001} = 1,000,000$$

10. If $E[X] < 0$ and $\theta \neq 0$ is such that $E[e^{\theta X}] = 1$, prove that $\theta > 0$.

Solution:

The function $f(x) = e^{\theta x}$ is convex. Jensen's inequality states that

$$E[f(X)] \geq f(E[X])$$

$$E[e^{\theta X}] \geq e^{\theta E[X]}$$

$$1 \geq e^{\theta E[X]}$$

$$0 > \theta E[X]$$

$$0 < \theta$$