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.

   3.6 $\{B_i\}_{i\in I}$, $B_i \in \mathscr{F}$, I finite or countable, such that $\cup_{i\in I} B_i = \Omega$ and $B_i \cap B_j = \emptyset \ \forall i \neq j$. From the last assumption, note that $A \cap B_i \subset B_i$, and hence $(A \cap B_i) \cup (A \cap B_i) = \emptyset \ \forall i \neq j$. Since the B_i form a partition of Ω,

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

   Now by Def 3.1.5 (ii), since the $A \cap B_i$ are pairwise disjoint,

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

   which is what we were asked to show. $\square$

   3.8 If $\{E_k\}_{k=1}^n$ is a collection of independent events, then prove that

   $$P(\cup_{k=1}^n E_k) = 1 - \prod_{k=1}^n (1 - P(E_k)).$$

   By definition of a independent collection of n events, for any m in the collection,

   $$P(\cap_{k=1}^m B_{i_k}) = \prod_{k=1}^m P(B_{i_k})$$

   And by De Morgan's laws, for any indexing set I, the following hold:

   $$\overline{\bigcap_{i\in I} A_i} \equiv \bigcup_{i\in I} \overline{A_i}, \quad \text{and} \quad \overline{\bigcup_{i\in I} A_i} \equiv \bigcap_{i\in I} \overline{A_i},$$

   Further, we can show that the finite collection of complements of events from an independent collection is itself independent.

   $$P(\overline{D} \cap C) = P(C \setminus (D \cap C)) = P(C) - P(D \cap C)$$

   $$= P(C) - P(D) \cdot P(C) = P(C) \cdot (1 - P(D)) = P(C) \cdot P(\overline{D})$$

   Taking $D = E_1$, it follows by induction on k that all sub-collections of $\overline{E_1} \cup \{E_k\}_{k=2}^n$ are independent (the ones not containing $\overline{E_1}$ are trivially independent by our assumptions). Since the collection $\overline{E_1} \cup \{E_k\}_{k=2}^n$ is not structurally

different from $\{E_k\}_{k=1}^n$, it follows by replacing every event by its complement in a finite number of n steps, that $\{\overline{E_k}\}_{k=1}^n$ is independent. The proof then follows in one line:

$$1 - P(\cup_{k=1}^n E_k) = P(\overline{\cup_{k=1}^n E_k}) = P(\cap_{k=1}^n \overline{E_k}) = \prod_{k=1}^n P(\overline{E_k}) = \prod_{k=1}^n (1 - P(E_k))$$

$\square$

3.11 From problem: #people in database $= 1mill$; #people who could commit crime $= 250mill$; $P(\text{s tests } +) = \frac{1}{3mill}$

With these given, and Karl's assurance that we could take $P(\text{s tested } +|\text{s} = \text{crime})$ as the ratio of #people in database and #people who could commit crime, we can use Bayes' Rule to find:

$$P(\text{s} = \text{crime}|\text{s tested } +) = \frac{P((\text{s tested } +) \cap (\text{s} = \text{crime}))}{P(\text{s tested } +)}$$

$$= \frac{P(\text{s tested } +|\text{s} = \text{crime})P(\text{s} = \text{crime})}{P(\text{s tested } +)} = \frac{\frac{1mill}{250mill}\frac{1}{250mill}}{\frac{1}{3mill}} = \frac{3}{62500} \approx 0.000048$$

$\square$

3.12 Probabilities in the Monty Hall Problem The special case is implied by the general one. I will label the doors as 1 to n. By symmetry, we assume without loss of generality that door #1 is chosen. Initially, the chance of the car being behind any one of them is $P(d) = \frac{1}{n}, \forall d \in \{1, ..., n\}$. Now the host opens n-2 doors, again, without loss of generality doors #3 to #n. Now observe that since # 1 is chosen, $P(\#1 opened) = 0$, and only one of the remaining $n - 1$ doors will stay open. Hence $P(\text{d closed}) = \frac{1}{n-1}$. Likewise, we know that the car must be behind #1 or #2. $P(\#2 \text{ closed}| \text{ car} = \#2) = 1$, since the door could not be opened without revealing the car, and thus forces the host to uncover the other doors,, and $P(\#2 \text{ closed}|\text{car} = \#1) = \frac{1}{n-1}$, since any other of the $n - 1$ door combinations could have been chosen. Thus,

$$P(\text{car} = \#1|\#2 \text{ closed}) = \frac{P(\#2 \text{ closed}|\text{car} = \#1)P(\text{car} = \#1)}{P(\#2 \text{ closed})} = \frac{1}{n}$$

and

$$P(\text{car} = \#2|\#2 \text{ closed}) = \frac{P(\#2 \text{ closed}|\text{car} = \#2)P(\text{car} = \#2)}{P(\#2 \text{ closed})} = \frac{n - 1}{n}.$$

Hence, for $n = 3$, switching results in success with probability $\frac{2}{3}$, and for $n = 10$, switching results in success with probability $\frac{9}{10}$. $\square$

3.16 Prove Theorem 3.3.18 The theorem is the following: Let X be a random variable with $E[X] = \mu$. Denote by X^2 the random variable given by $X^2(\omega) = (X(\omega))^2$. Then $Var[X] = E[X^2] - E[X]^2 = [X^2] - \mu^2$.

proof $Var[X] = E[(X - \mu)^2] = E[X^2 - 2\mu X + \mu^2] = E[X^2] - 2\mu E[X] + \mu^2$
$= E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - \mu^2 \ \square$

3.33 Let $B = B(n, p)$ be a random binomial variable for n trials with parameter p. Prove for any $\epsilon > 0$ that

$$P(|\frac{B}{n} - p| \geq \epsilon) \leq \frac{p(1-p)}{n\epsilon^2}$$

(proof) Define $X = \frac{B}{n}$ as a new random variable. Observer that $E[X] = \frac{np}{n} = p$ and $Var[X] = Var[\frac{B}{n}] = \frac{Var[B]}{n^2} = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$. Using Chebhyshev's inequality with $X = \frac{B}{n}$ and $\epsilon =$, the statement trivially follows:

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2} \leftrightarrow P(|X - p| \geq \epsilon) \leq \frac{p(1-p)}{n\epsilon^2} \quad \square$$

3.36 From problem:

The central limit theorem states that $P(\frac{S_n - n\mu}{\sigma \sqrt[2]{n}} \leq \gamma) \to \Phi(\gamma)$ where the latter is the standard normal distribution cdf. $P(X \leq 5500) = \Phi(\frac{5500 - 5000}{\sqrt{6242}\sqrt{0.801 \cdot 0.199}}) = \Phi(15.85) \approx 1$, so there is virtually no danger of more than 5500 students. $\square$

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 $P(A \cap B \cap C) \neq P(A)P(B)P(C)$.

      Solution Let A, B, C be the events that in flipping a coin, the first and second, first and third, and second and third flips have the same return, respectively. Note that the probability of each event is $\frac{1}{2}$, and for two events to occur, all flips have to have the same return, which happens with probability $\frac{1}{4}$. Hence, $P(E \cap D) = \frac{1}{4} = (\frac{1}{2})^2 = P(E)P(D)$ for $E \neq D$ Hence the events are pairwise independent. However, $A \cap B \cap C$ is also the event that all flips are the same, hence the collection of the three events is not independent: $P(A \cap B \cap C) = \frac{1}{4} \neq (\frac{1}{2})^3 = P(A)P(B)P(C)$. $\square$

   (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.)

Solution Let A, B be the events that in flipping a coin, the first and second, first and third have the same return, respectively. Let C be the event that either the first two flips are heads, or that the flips are alternating (i.e. HTH or THT). All individual events thus have probability $\frac{1}{2}$. Then $A \cap B$ is as in (a) independent, $A \cap C = \{HHH, HHT\}$, and $A \cap B \cap C = \{HHH\}$. Thus, $P(A \cap B) = \frac{1}{4} = P(A)P(B)$, $P(A \cap C) = \frac{1}{4} = P(A)P(C)$, $P(A \cap B \cap C) = \frac{1}{8} = P(A)P(B)P(C)$. However, $B \cap C = \{HHH, HTH, THT\}$ so $P(B \cap C) = \frac{3}{8} \neq \frac{1}{4} = P(B)P(C)$, as desired. $\square$

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

Solution To show that Benford's Law is a well-defined discrete probability distribution, we check:

- $0 \leq P(d) \leq 1$

$$P(d) = log_{10}(1 + \frac{1}{d}) > 0 \quad \text{since} \quad \frac{1}{d} > 0, \quad \text{and}$$

$$P(d) = log_{10}(1 + \frac{1}{d}) < 1 \quad \text{since} \quad 1 + \frac{1}{d} < 10$$

- $P(\Omega) = \sum_{d=1}^{9} P(d) = 1$

$$P(\Omega) = \sum_{d=1}^{9} P(d) = \sum_{d=1}^{9}(log_{10}(1 + \frac{1}{d}) = \sum_{d=1}^{9}(log_{10}(\frac{d+1}{d})$$

$$= \sum_{d=1}^{9}(log_{10}(d+1) - log_{10}(d) = log_{10}(10) - log_{10}(1) = 1$$

- Additivity: each of the d are unique (and thus pairwise disjoint), and

$$P(\cup_{d \in D \subset \Omega} d) = \sum_{d \in D \subset \Omega} P(d)$$

This shows that Benford's Law is well-defined. $\square$

4. A person tosses a fair coin until a tail appears for the first time. If the tail appears on the nth flip, the person wins 2^n dollars. Let the random variable X denote the player's winnings.

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

   Proof Note that the probability of a player being able to flip the coin is $(\frac{1}{2})^{n-1}$ (that is, the player must have tossed a head in all $n - 1$ previous rounds. Thus, the probability of getting a tail on the nth flip is $(\frac{1}{2})^n$, and the payoff is 2^n. But then $E[X] = \sum_{n=1}^{\infty}(\frac{1}{2})^n 2^n = \sum_{n=1}^{\infty} 1 = +\infty$

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

Proof $E[ln(X)] = \sum_{n=1}^{\infty} (\frac{1}{2})^n ln(2^n) = \sum_{n=1}^{\infty} (\frac{n}{2^n}) ln(2) = 2ln(2)$ where we use that

$$S = \sum_{n=1}^{\infty} (\frac{n}{2^n}) = S - \frac{1}{2}S + \frac{1}{2}S = \sum_{n=1}^{\infty} (\frac{n}{2^n}) - \frac{1}{2}\sum_{n=1}^{\infty} (\frac{n}{2^n}) + \frac{1}{2}S$$

$$= \sum_{n=1}^{\infty} (\frac{n}{2^n}) - \sum_{n=1}^{\infty} (\frac{n}{2^{n+1}}) + \frac{1}{2}S = \sum_{n=1}^{\infty} (\frac{n}{2^n}) - \sum_{n=2}^{\infty} (\frac{n-1}{2^n}) + \frac{1}{2}S$$

$$= \sum_{n=1}^{\infty} (\frac{n}{2^n}) - \sum_{n=1}^{\infty} (\frac{n-1}{2^n}) + \frac{1}{2}S = \sum_{n=1}^{\infty} (\frac{1}{2^n}) + \frac{1}{2}S = 1 + \frac{1}{2}S \Rightarrow S = 2$$

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?

The investors should invest in foreign currency.

Proof By symmetry, we look at the U.S. investor: she could retain her wealth in her own currency (USD), and with certainty have the $(1 + r)$ the amount of USD in the future. Now observe that if she invests in foreign currency, she will have her initial wealth in CHF, and thus, depending on the state, either 1.25 or 0.8 USD per CHF in the future. Here note, that

E[wealth in USD | initwealth traded for CHF] $= (1+r)\frac{1}{2}(1.25+0.8)(initwealth) = 1.025*(1+r)$ initial wealth. Thus, investing all the wealth in foreign currency is the best strategy, assuming the problem's conditions hold. $\square$

6. Consider a probability measure space with $\Omega = [0, 1]$.

   (a) Construct a random variable X such that $E[X] < \infty$ but $E[X^2] = \infty$.

   Example A t_2 distribution has finite mean and infinite variance, i.e. infinite square mean. $\square$

   (b) Construct random variables X and Y such that $P(X > Y) > \frac{1}{2}$ but $E[X] < E[Y]$.

Example Let X be a well-defined random variable. Let $Y(\omega) = X(\omega) + 5\epsilon$ with probability $\frac{1}{5}$ and $Y(\omega) = X(\omega) - \epsilon$ with probability $\frac{4}{5}$. Then $P(X > Y) = \frac{4}{5} > \frac{1}{2}$, and $E[X] < E[Y] = E[\frac{1}{5}(X + 5\epsilon) + \frac{4}{5}(X - \epsilon)] = E[X] + \frac{1}{5}\epsilon.\square$

(c) Construct random variables X, Y, and Z such that
$P(X > Y)P(Y > Z)P(X > Z) > 0$ and $E(X) = E(Y) = E(Z) = 0$.

Example If $P(X > Y)P(Y > Z)P(X > Z) > 0$ is interpreted as a product, we can simply take all to be independent normal distributions centered at 0.

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) $Y \sim N(0,1)$.

True $P(Y) = P(XZ) = P(X|Z = -1) \cdot P(Z = -1) + P(X|Z = 1) \cdot P(Z = 1) = P(X)$ hence Y follows the same distribution as X.

(b) $P(|X| = |Y|) = 1$.

True This follows from (a). $(|Y| = |XZ| = |X||Z| = |X|)$

(c) X and Y are not independent.

True $P(X > \epsilon \cap Y > \epsilon) = P(Z = 1)P(X > \epsilon) = \frac{1}{2}P(X > \epsilon) \neq P(X > \epsilon) \cdot P(Y > \epsilon)$

(d) $Cov[X, Y] = 0$.

True By definition, $Cov[X, Y] = E[XY] - E[X]E[Y] = E[X^2 \cdot Z] = 0$ by independence of X and Z.

(e) If X and Y are normally distributed random variables with $Cov[X, Y] = 0$, then X and Y must be dependent.

False Just because this applies for this particular example, it must not hold for all. Take two independent normal random variables. Their covariance must be 0 since E[XY]=E[X]E[Y].

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

m $P(m \leq x) = 1 - (P(X > x))^n = 1 - (1-x)^n$. Thus the pdf is $p(m) = n \cdot (1-x)^{n-1}$, and the expectation is

$$E[m] = \int_0^1 xp(x)dx = n \cdot \int_0^1 x(1-x)^{n-1}dx = n \cdot \int_0^1 -(1-x)^n + (1-x)^{n-1}dx$$

$$= n \cdot \left(-\frac{1}{n+1} + \frac{1}{n}\right) = \frac{1}{n+1}$$

M $P(M \leq x) = (P(X \leq x))^n = x^n$. Thus the pdf is $p(M) = n \cdot x^{n-1}$, and the expectation is

$$E[M] = \int_0^1 xp(x)dx = n \cdot \int_0^1 x \cdot x^{n-1}dx = n \cdot \int_0^1 x^n dx = \frac{n}{n+1}$$

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%?

   Answer 2% of 500 are 10, so we are looking for $P(490 \leq X \leq 510)$, where the random variable X denotes the number of good states in 1000 periods. $P(490 \leq X \leq 510) = \sum_{X=490}^{510} P(X) = \sum_{X=490}^{510} \binom{1000}{X}(\frac{1}{2})^{1000} = 0.49334$ by Wolfram Alpha. Note that for Using Central Limit Theorem, this approximately follows a normal distribution with mean 500 and variance of 250. Then

   $$P(490 \leq X \leq 510) = P(X \leq 510) - P(X < 490) = P(X \leq 510) - (1 - P(X \geq 490))$$

   $$= 2P(X \leq 510) - 1 = 2\Phi\left(\frac{510-500}{\sqrt{250}} \approx 0.63\right) - 1 = 2 \cdot 0.7357 - 1 = 0.4714$$

   (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%?

   Answer To ensure that $P(0.495 \leq P(goodstate) \leq 0.505)$, we consider Chebhyshev with X being the proportion of good states:
   $$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2} \leftrightarrow P(|X - 0.5| \geq 0.005) \leq \frac{(\frac{0.5}{\sqrt{n}})^2}{0.005^2} \quad \text{with} \quad \frac{(\frac{0.5}{\sqrt{n}})^2}{0.005^2} \leq 0.01$$
   To make the inequality bind (i.e to find the lowest n for which the inequality must hold, we find $n \geq 1000000$.

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

Proof   Using Jensen's inequality with $f(x) = -ln(x)$, we have that

$$0 = -ln(1) = -ln(E[e^{\theta X}]) \leq E[-ln(e^{\theta X})] = -E[\theta X] = -\theta E[X] = \theta|E[X]|$$

Hence, $\theta \geq 0$, and since $\theta \neq 0$, $\theta > 0$.