OSM Lab 2017: Math Pset 1

Wei Han Chia

Due: 26 June 2017

Problems from the Book

3.6

Let $(\Omega, \mathbf{F}, P)$ be a probability space. Let $\{B_i\}_{i \in I}$ be a collection of elements of $\mathbf{F}$ indexed by a finite/countable set I, such that $\Omega = \cup_{i \in I} B_i$ and $B_i \cap B_j = \emptyset$ for all $i \neq j$.

Consider any $A \in \mathbf{F}$. Now by finite additivity, $\sum_{i \in I} P(A \cap B_i) = P(\cup_{i \in I}(A \cap B_i))$. Now we want to show $\cup_{i \in I}(A \cap B_i) = A$.

Consider any $a \in A$. Now since $\cup_{i \in I} B_i = \Omega$, it follows that $a \in A \cap B_i$ for some i. Therefore $A \subset \cup_{i \in I} A \cap B_i$. Now consider any $a \in \cup_{i \in I} A \cap B_i$. Then $a \in A \cap B_i \subset A$, so we have shown that $\cup_{i \in I} A \cap B_i \subset A$. So we have proven that $\cup_{i \in I}(A \cap B_i) = A$.

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

3.8

Let $\{E_1, E_2, ..., E_n\}$ be a collection of independent events. We want to prove:

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

Lets denote $\cup_{k=1}^n E_k = B$. Then we know from our properties of probability spaces that $P(B) = 1 - P(B^c)$. We claim that $B^c = \cap_{k=1}^n E_k^c$. Consider $b \in B^c$. Then $b \notin \cup_{k=1}^n E_k \implies b \in E_k^c$ for all k, i.e. $b \in \cap_{k=1}^n E_k^c$. So $B^c \subset \cap_{k=1}^n E_k^c$.

Now consider $b \in \cap_{k=1}^n E_k^c$. Then $b \notin E_k \forall k$. It follows that $b \in (\cup_{k=1}^n E_k)^c$. So $\cap_{k=1}^n E_k^c \subset B^c$.

Now we have shown that $B^c = \cap_{k=1}^n E_k^c$. Now since we know that each E_k is independent, $P(B^c) = \pi_{k=1}^n P(E_k^c) = \pi_{k=1}^n (1 - P(E_k))$.

3.11

As per the hint, we want to first find $P(s = crime|stested+)$. We can find this using Bayes Rule.

$$P(s = crime|stested+) = \frac{P(stested + |s = crime)P(crime)}{P(stested+)} = \frac{\frac{1}{250} \frac{1}{250,000,000}}{\frac{1}{3,000,000}} = \frac{3}{250^2}$$

Therefore, we note that the probability that the match in the database was the one at the crime scene is $\frac{3}{250^2}$.

3.12

Let us consider first the 3 door problem. Now we know that the probability that the car is found behind any door (i) is $P(C_i) = \frac{1}{3}$. We also have the events of the host opening the door (D_i). Without loss of generality, say we pick door 1. Now consider if the host opens door 3. We know that $P(D_3|C_1) = \frac{1}{2}$, $P(D_3|C_2) = 1$, $P(D_3|C_3) = 0$, since the host would never open a door with the car behind it. We also know that $P(D_3) = P(D_2) = \frac{1}{2}$, since $P(D_1) = 0$.

Now using Bayes rule, $P(C_1|D_3) = \frac{P(D_3|C_1)P(C_1)}{P(D_3)} = \frac{1}{3}$. $P(C_2|D_3) = \frac{P(D_3|C_2)P(C_2)}{P(D_3)} = \frac{2}{3}$. So given that hte host opens one door, it is always better to switch as the conditional probability of that door containing the car would be higher.

We can generalize this to the case with n doors. In particular, we note that the probability of the other door containing the car given all other doors were opened would be $\frac{n-1}{n}$, while the probability of the original door containing the car given all the other doors were opened would be $1n$.

3.16

We want to prove $Var[X] = \mathbf{E}[X^2] - (\mu)^2$.

$$
\begin{aligned}
Var[X] &= \mathbf{E}[(X - \mu)^2] \\
&= \mathbf{E}[X^2 - 2\mu X + \mu^2] \\
&= \mathbf{E}[X^2] - 2\mu\mathbf{E}[X] + \mu^2 \quad \text{By linearity of Expectation Operator} \\
&= \mathbf{E}[X^2] - 2\mu^2 + \mu^2 \\
&= \mathbf{E}[X^2] - \mu^2
\end{aligned}
$$

3.33

Let $B = B(n, p)$. We want to apply Chebyshev's Inequality. $P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}$. To do this, consider $X = \frac{B}{n}$.

Now $\text{Var}[X] = \frac{\text{Var}[B]}{n^2} = \frac{p(1-p)}{n}$. Also, $\mathbf{E}[X] = p$. Then, we can simply apply Chebyshev's Inequality to X to conclude:

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

3.36

Now lets assume our university admits 6242 students. Then we know that the number of students is a sum of independent Bernoulli(0.801) random variables. Using the CLT, we know that the number of admitted students follows approximately a $N(5000, 5000(0.199))$ distribution.

Now $P(Students > 5500) = 1 - P(Students \leq 5500) \approx 0$.

Problem 2

(a)

Consider the set of equally possible points $\{a_1, a_2, a_3, a_4\}$. Now let us define the following events

$$
\begin{aligned}
A &= (a_1, a_2) \\
B &= (a_2, a_3) \\
C &= (a_3, a_1)
\end{aligned}
$$

Now we note that for any $X, Y \in \{A, B, C\} X \neq Y$, $P(X \cap Y) = \frac{1}{4} = P(X)P(Y)$. However, we note that $P(A \cap B \cap C) = 0 \neq P(A)P(B)P(C)$.

(b)

Consider the set of equally possible points $\{a_1, a_2, ..., a_8\}$. Now let us define the following events

$$
\begin{aligned}
A &= (a_1, a_2, a_3, a_4) \\
B &= (a_1, a_2, a_5, a_6) \\
C &= (a_1, a_3, a_7, a_8)
\end{aligned}
$$

Now $P(A) = P(B) = P(C) = \frac{1}{2}$. Also, $P(A \cap B) = \frac{1}{4} = P(A \cap C)$. In addition, $P(A \cap B \cap C) = \frac{1}{8} = P(A)P(B)P(C)$. However, note that $P(B \cap C) = \frac{1}{8} \neq P(B)P(C) = \frac{1}{4}$.

Problem 3

From class, we defined a set of numbers to satisfy Benford's law if $P(d) = \log(1 + \frac{1}{d})$ where d is the leading digit of the set of numbers. Note that $P(d) = \log(\frac{d+1}{d}) = \log(d + 1) - \log(d)$.

Now we want to show that this satisfies a probability distribution. In essence, we want to show that $\sum_{d=1}^{9} P(d) = 1$. This follows from evaluating the sum, and noting that this reduces to $\log(10) - \log(1) = 1$. So the discrete distribution defined by Benford's Law is indeed a valid probability distribution.

Problem 4

Suppose a person tosses a fair coin until a tail appears the first time. They get 2^n if the tail appears on the $n-thflip$. Let X denote the player's winnings.

(a)

$$\mathbf{E}[X] = \sum_{i=1}^{\infty} 2^i (\frac{1}{2})^i = \sum_{i=1}^{\infty} 1 = +\infty$$

(b)

Now let us consider when the agent has log utility.

$$\mathbf{E}[\ln X] = \sum_{i=1}^{\infty} (\frac{1}{2})^i i \ln 2 = \ln 2 \sum_{i=1}^{\infty} \frac{i}{2^i}$$

$$\frac{1}{2} \sum_{i=1}^{\infty} \frac{i}{2^i} = \sum_{i=1}^{\infty} \frac{i}{2^i} - \sum_{i=1}^{\infty} \frac{i}{2^{i+1}} = \sum_{i=1}^{\infty} \frac{1}{2^i} = 1$$

$$\implies \mathbf{E}[\ln X] = 2 \ln 2$$

Problem 5

Consider the problem of 2 investors, one in Switzerland, and one in the US, where the exchange rate is currently 1:1. Also consider a fixed interest rate of (r). We note that in the future, we have an exchange rate of 1.25:1 with probability 0.5, and an exchange rate of 1:1.25 with probability 0.5.

Without loss of generality, consider the problem of the US investor.

$$\mathbf{E}[\text{returns}] = \begin{cases} 1 + r & \text{investing in USD} \\ (1+r)[0.5(1.25) + 0.5(1/1.25)] = (1+r)(1.025) & \text{investing in CHF} \end{cases}$$

So in this case, despite the potential for a risk free investment in local currency, the US investor would have a higher expected return by investing in CHF. Similarly, a Swiss investor would have a higher expected return by investing in USD.

Problem 6

(a)

Consider the pareto distribution with scale factor 1, $\alpha = 32$. Then, the pdf of this distribution is $f(x) = \frac{(3/2)^{3/2}}{x^{5/2}}$. We also know that the mean of this distribution is 3, and the variance is ∞.

This is an example of a fat tailed distribution.

(b)

Let $Y \sim N(0,1)$. We define X as follows:

$$X = \begin{cases} Y + 1 & \text{with probability 2/3} \\ Y - 3 & \text{with probabiity 1/3} \end{cases}$$

Now clearly $P(X > Y) = \frac{2}{3} > \frac{1}{2}$. However, note that $\mathbf{E}[X] = \mathbf{E}[Y] + 2/3 - 1 < \mathbf{E}[Y]$.

(c)

We want a set of random variables such that $P(X > Y)P(Y > Z)P(X > Z) > 0$, but $\mathbf{E}[X] = \mathbf{E}[Y] = \mathbf{E}[Z] = 0$. Consider X, Y, Z independently distributed $\sim N(0,1)$. Then we have $\mathbf{E}[X] = \mathbf{E}[Y] = \mathbf{E}[Z] = 0$.

Further note that $P(X > Y) = P(X - Y > 0) = \frac{1}{2}$. This holds for all three probabilities, such that $P(X > Y)P(Y > Z)P(X > Z) = \frac{1}{8} > 0$.

Problem 7

Let the random variables X and Z be independent with $X \sim N(0,1)$, $P(Z = 1) = P(Z = -1) = \frac{1}{2}$. Define $Y = XZ$.

(a)

Lets consider the cdf of Y.

$$P(Y < y) = P(XZ < y) = P(X < y|Z = 1).P(Z = 1) + P(-X > -y|Z = -1).P(Z = -1)$$
$$= P(X < y) \quad \text{Since the normal distribution is symmetric}$$

So $Y \sim N(0,1)$ as well.

(b)

Consider $|Y|$.

$$|Y| = |XZ| = |X||Z| = |X|$$

So it follows that $P(|X| = |Y|) = P(|X| = |X|) = 1$.

(c)

Consider $P(Y < y \cap X < x) = 0.5P(X < -1) \neq P(Y < -1)P(X < -1)$. So the two events are not independent.

(d)

$$Cov(X,Y) = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y]$$
$$= \mathbf{E}[X^2 Z] = \mathbf{E}[X^2]\mathbf{E}[Z]$$
$$= 0$$

So the two events have 0 covariance.

(e)

Combining the results from (c) and (d) show us a pair of normally distributed random variables which have 0 covariance, but are not independent.

Note that this statement only works for a pair of **jointly** normally distributed random variables.

Problem 8

Consider $m = \min\{X_1, ...X_n\}$, $M = \max\{X_1, ..., X_n\}$. Where each $X_i \sim Uniform(0,1)$. We first derive the cdf, then the pdf through differentiation, and finally the expected values of m and M.

$$F_m(x) = P(m \leq x) = 1 - (1 - P(X < x))^n$$
$$= 1 - (1 - x)^n$$
$$f_m(x) = n(1 - x)^{n-1}$$
$$\mathbf{E}(m) = n \int_0^1 x(1 - x)^{n-1} dx$$
$$= \frac{1}{n+1} \quad \text{The integral is a beta function with parameters 2, n}$$

$$F_M(x) = P(M \leq x) = (P(X, x))^n$$
$$= x^n$$
$$f_M(x) = nx^{n-1}$$
$$\mathbf{E}(M) = n \int_0^1 x^n dx$$
$$= \frac{n}{n+1}$$

Problem 9

We want to simulate a dynamic economy with 2 states, and a probability of shock of 0.5. It is clear that given a fixed number of periods n, the distribution of good states (X) is given by a binomial distribution $B(n, 0.5)$.

(a)

Consider $n = 1000$. From the CLT, we know that $X \sim N(500, 250)$. We want to find the probability that the number of good states differs by at most 2%, i.e. that $490 < X < 510$. Now we can use our normal distribution to get:

$$P(X \text{ differs from } 500 \text{ by } 2\%) = 0.4732$$

(b)

Now let Y be the proportion of good states. By the central limit theorem we know that $Y \sim N(\frac{1}{2}, \frac{1}{4n})$. Now we want $P(|Y - 0.5| \leq 0.005)$. From our Chebyshev's inequality, we can get an upper bound for $P(|Y - 0.5| \geq 0.005)$.

$$P(|Y - 0.5| \geq 0.005) \leq \frac{1}{4(0.005)^2 n}$$

In order for $P(|Y - 0.5| \leq 0.005) \geq 0.99$, we need $P(|Y - 0.5| \geq 0.005) \leq 0.01$, i.e. $\frac{1}{4(0.005)^2 n} \leq 0.01$. Solving this inequality gives $n \geq 1000000$.

Problem 10

We know that $\mathbf{E}[X] < 0$ and $\theta \neq 0$ such that $\mathbf{E}[e^{\theta x}] = 1$. We want to prove that $\theta > 0$.

Proof. For contradiction, assume that $\theta < 0$. Now we know that $e^{\theta x}$ is a convex function, so $e^{\theta \mathbf{E}[X]} \leq \mathbf{E}[e^{\theta x}] = 1$. However, since we know that $\mathbf{E}[X] < 0$, and $\theta < 0$, it follows that $\theta \mathbf{E}[X] > 0$. Since e is monotone increasing, and $e^0 = 1$, it must be that $e^{\theta \mathbf{E}[X]} > 1$, a contradiction.
   Therefore $\theta > 0$. $\qquad\square$