Math, Problem Set #1, Probability Theory

Francesco Furno

Joint work with Ildebrando Magnani

OSM Lab, Karl Schmedders

Due Monday, June 26 at 8:00am

Exercise 1

Exercise 3.6

Notice that $A = A \cap \Omega = A \cap (\cup_i B_i) = \cup_i (A \cap B_i)$. Notice that for every $i \neq j$, $(A \cap B_i) \cap (A \cap B_j) = A \cap B_i \cap B_j = \emptyset$. Therefore, we have that $P(A) = P(\cup_i (A \cap B_i)) = \sum_i P(A \cap B_i)$, where the last equality follows from additivity of P.

Exercise 3.8

By DeMorgan's law we have that:

$$P(\cup_k E_k) = (P(\cap_k E_k^c)^c)$$

where E^c is the complement of E. Therefore we have that:

$$\begin{aligned} P(\cup_k E_k) &= 1 - P(\cap_k E_k^c) \\ &= 1 - \prod_k P(E_k^c) \\ &= 1 - \prod_k (1 - P(E_k)) \end{aligned}$$

where the second equality follows from independence of the collection $\{E_k\}$.

Exercise 3.11

Applying Bayes' Rule we learn that:

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

From the information provided we know that $P(s\ tested+) = 1$ and that experts say that $P(stested + | s = crime) = (3 \times 10^6)^{-1}$. The size of the population of people old enough to have committed the crime helps us defining $P(s = crime) = (250 \times 10^6)^{-1}$. Therefore:

$$P(s = crime | s\ tested+) = (3 \times 10^6)^{-1} \times (250 \times 10^6)^{-1} = \frac{1}{750 \times 10^{12}}$$

Exercise 3.12

Call the probability of finding the car at the first attempt $P(car)$. Naturally, we have that $P(car) = 1/3$. Notice that the probability of the car being in one of the two other remaining doors is $P(notcar) = 2/3$.
After Monty opens one of the other two doors, the probability of finding the car by sticking to your first choice remains the same and equals 1/3. However, if you switch to the remaining door, the probability of finding the car increases. There are two cases: either the car was in the first door that you chose, and in this case you lose for sure, or the car was in one of the other two, and in this case you find it for sure (thanks to Monty who ruled out the door with the goat). Therefore:

$$\begin{aligned} P(car\ after\ switch) &= P(car) \times P(find|car) + P(notcar) \times P(find|notcar) \\ &= 1/3 \times 0 + 2/3 \times 1 = 2/3 \end{aligned}$$

which is clearly greater than 1/3.
As long as Monty opens out all but one of the remaining doors, the general principle does not change: it is better to switch to the remaining door. With 10 doors the result is even more apparent: the probability of finding the car with the first choice is now $P(car) = 1/10$, and $P(notcar) = 9/10$. Thanks to Monty, if the car is in one of the remaining 9 doors (which happens wp 9/10), you will find it for sure. Thus:

$$P(car\ after\ switch) = 9/10$$

Exercise 3.16

Applying the definition of variance of a random variables, solving the square and exploting the fact that the expectation operator is linear we have that:

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

Exercise 3.33

Let $\tilde{B} = B/n$. Notice that $E[\tilde{B}] = E[B]/n$ and $V[\tilde{B}] = V[B]/n^2$, where $E[B] = np$ and $V[B] = np(1-p)$. Therefore, $E[\tilde{B}] = p$ and $V[\tilde{B}] = p(1-p)/n$. Apply the Chebyschev's Inequality on $\tilde{B}$:

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

Exercise 3.36

Let X_i be identical and independently distributed bernoulli random variables with paramenter $p = 0.801$, with $i = 1, 2, ..., n$, where $n = 6242$. Wlog, define:

$$X_i = \begin{cases} 1 & if\ student\ accepts\ the\ offer \\ 0 & if\ student\ rejects\ the\ offer \end{cases}$$

We are interested in $S_n = \sum_i X_i$, because it represents the total number of students accepting the offer. Since the random variables X_i are i.i.d., the CLT tells us that:

$$\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} = \frac{1}{\sqrt{n}}\frac{S_n - n\mu}{\sigma} \xrightarrow{D} N(0,1)$$

where $\bar{X}_n = \frac{\sum_i X_i}{n}$. Notice that $S_n = \bar{X}_n \times n$ and that we are interested in $P(S_n > 5500)$. Notice that all the following events have identical measure:

$$\begin{aligned} S_n > 5500 &= S_n - n\mu > 5500 - n\mu \\ &= \frac{S_n - n\mu}{\sigma} > \frac{5500 - n\mu}{\sigma} \\ &= \frac{1}{\sqrt{n}}\frac{S_n - n\mu}{\sigma} > \frac{1}{\sqrt{n}}\frac{5500 - n\mu}{\sigma} \end{aligned}$$

where $\mu = E[X_i] = p = 0.801$, so that $n\mu = np = 6242 \times 0.801 = 5000$, and $\sigma = V[X_i] = p(1-p) = 0.801 \times 0.199 = 0.1594$. Therefore :

$$\begin{aligned} P(S_n > 5500) &= P(\frac{1}{\sqrt{n}}\frac{S_n - n\mu}{\sigma} > \frac{1}{\sqrt{n}}\frac{5500 - n\mu}{\sigma}) \\ &= P(Z_n > \frac{1}{\sqrt{n}}\frac{5500 - n\mu}{\sigma}) \\ &= P(Z_n > \frac{1}{\sqrt{6242}}\frac{5500 - 5000}{0.1594}) \\ &= P(Z_n > \frac{3136.76}{79}) \\ &= P(Z_n > 39.71) \end{aligned}$$

The CLT tells us that Z_n is approximately distributed as a standard normal, so that:

$$P(S_n > 5500) \approx P(Z_n > 39.71) = 1 - \Phi(39.71) \approx 0$$

where $\Phi(.)$ is the CDF of a standard normal random variable. Put it differently, given the assumptions, it is almost impossible that more than 5500 students will accept the University's offer.

Exercise 2

Part a

Perform two independent tosses of a coin. Let A be the event that a head shows up in the first toss, B the event that a head shows up in the second toss, and C the event that precisely one head shows up either in the first or second toss. Notice that:

$$\begin{aligned}
P(A) &= \frac{1}{2} \\
P(B) &= \frac{1}{2} \\
P(C) &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \\
P(A \cap B) &= \frac{1}{4} \\
P(A \cap C) &= P(H, T) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \\
P(B \cap C) &= P(T, H) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}
\end{aligned}$$

So that $P(A \cap B) = P(A) \times P(B)$, $P(A \cap C) = P(A) \times P(C)$ and $P(B \cap C) = P(B) \times P(C)$. However:

$$P(A \cap B \cap C) = 0 \neq P(A) \times P(B) \times P(C) = \frac{1}{8}$$

Part b

Perform three independent tosses of a coin. Consider the following three events:

1. $A =$ head at first toss;
2. $B =$ precisely one head, either at first or third toss;
3. $C =$ precisely one head, either at first or second toss.

Notice that $P(A) = 1/2$, $P(B) = 1/4$ and $P(C) = 1/4$. We have that:

$$\begin{aligned}
P(A \cap B) &= \frac{1}{4} = P(A) \times P(B) \\
P(A \cap C) &= \frac{1}{4} = P(A) \times P(C) \\
P(B \cap C) &= \frac{1}{8} \neq P(B) \times P(C) \\
P(A \cap B \cap C) &= P(A) \times P(B) \times P(C) = \frac{1}{8}
\end{aligned}$$

Exercise 3

Let $\mathcal{D} = \{1, 2, ..., 9\}$ and $P(d) = log_{10}(1 + 1/d) = log_{10}(\frac{1+d}{d}) = log_{10}(1 + d) - log_{10}(d)$ be the probability distribution of the random variable d with support $\mathcal{D}$. To show that $P(.)$ is well-defined, we need to check that:

1. $P(\Omega) = 1$;

2. $P(.)$ is additive.

$P(.)$ is additive by assumption. Therefore, we need to show that $P(\Omega) = 1$. Notice that:

$$\Omega = \cup_{d=1}^{9}\{d\}$$

Therefore:

$$\begin{aligned} P(\Omega) &= \sum_{d=1}^{9} log_{10}(1+d) - log_{10}(d) \\ &= log_{10}(10) - log_{10}(1) \\ &= 1 - 0 = 1 \end{aligned}$$

Therefore, $P(.)$ is a well-defined probability distribution.

Exercise 4

Part a

Notice that the support of X is given by $\mathcal{X} = \{2, 2^2, 2^3, ...\}$ and that $P(X = 2^n) = \frac{1}{2^n}$. Therefore:

$$\begin{aligned} E[X] &= \sum_{n=1}^{+\infty} 2^n \times \frac{1}{2^n} \\ &= \sum_{n=1}^{+\infty} 1 = +\infty \end{aligned}$$

Part b

Now consider the new random variable $Y = ln(X)$, with support $\mathcal{Y} = \{ln(2), ln(2^2), ln(2^3), ...\}$. We have that:

$$\begin{aligned} E[Y] = E[ln(X)] &= \sum_{n=1}^{+\infty} ln(2^n) \times \frac{1}{2^n} \\ &= \sum_{n=1}^{+\infty} n \times ln(2) \times \frac{1}{2^n} \\ &= ln(2) \sum_{n=1}^{+\infty} \frac{n}{2^n} = 2 \times ln(2) \end{aligned}$$

Exercise 5

Wlog, consider the investor from the U.S.. He has two alternatives:

1. Invest in the U.S. and obtain a fixed (safe) return R;

2. Convert USD in CHF, invest in Switzerland, obtain a fixed (safe) return R and convert CHF back into USD. Obviously, this investment is subject to exchange rate uncertainty.

The return from the investment in Switzerland is given by the random variable R^{CHF}. Notice that:

$$R^{CHF} = \begin{cases} R \times (1.25) & \textit{with probability } \frac{1}{2} \\ R \times (1.25^{-1}) & \textit{with probability } \frac{1}{2} \end{cases}$$

The expected return of investing in Switzerland is therefore equal to:

$$E[R^{CHF}] = \frac{1}{2} \times 1.25 \times R + \frac{1}{2} \times 0.8 \times R = 1.025 \times R$$

which is greater than $R = R^{USA}$. Therefore, a risk-neutral U.S. investor that only considers expected returns will prefer to invest in Switzerland. Symmetrically, a Swiss investor will prefer to invest in the U.S.. Hence the paradox.

Exercise 6

Part a

Let $X \sim Pareto(x_m, \alpha)$, with $\alpha \in (1, 2]$. Then:

- $E[X] < +\infty$
- $V[X] = +\infty$

Exercise 7

Part a

Let $F(y)$ be the CDF of Y. Then we have that:

$$\begin{aligned} F(y) &= P(Y < y) \\ &= P(XZ < y) \\ &= P((X < y \cap Z = 1) \cup (X > -y \cap Z = -1)) \\ &= P((X < y \cap Z = 1)) + P((X > -y \cap Z = -1)) \\ &= \frac{1}{2} \times P(X < y) + \frac{1}{2} \times P(X > -y) \\ &= \frac{1}{2} \times P(X < y) + \frac{1}{2} \times P(X > -y) \\ &= \frac{1}{2} \times \Phi(y) + \frac{1}{2} \times (1 - \Phi(-y)) \end{aligned}$$

where $\Phi(.)$ is the CDF of a standard normal random variable. Given that $\Phi(y) = 1 - \Phi(-y)$ we have that:

$$F(y) = \Phi(y)$$

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

Part b

By definition of Y and Z we have:

$$|Y| = |ZX| = \begin{cases} |-X| & wp\ 1/2 \\ |X| & wp\ 1/2 \end{cases}$$

Notice that by definition of absolute value we have that $|X| = |-X|$, so that $|Y| = |X|$ with probability one.

Part c

Let $k \in \mathbb{R}$. Notice that:

$$-k < Y < k \Rightarrow -k < ZX < k \Rightarrow -k < X < k$$

Therefore the two random variables are not independent.

Part d

From the definition of covariance:

$$Cov[X, Y] = E[XY] - E[X]E[Y]$$

Given that $Y = ZX$, and that X and Z are independent, it follows that:

$$\begin{aligned} Cov[X, Y] &= E[XXZ] - E[X]E[XZ] \\ &= E[X^2]E[Z] - E[X]^2E[Z] \\ &= E[Z](E[X^2] - E[X]^2) = 0 \end{aligned}$$

since $E[Z] = \frac{1}{2} \times 1 + \frac{1}{2} \times (-1) = 0$.

Part e

Points (a) and (d) disprove the claim. In fact, the correct statement would be: *If X and Y are* ***jointly*** *normally distributed random variables with $Cov[X, Y] = 0$, then X and Y must be independent.* Our X and Y are marginally normally distributed, but not jointly.

Exercise 8

Let $F(x) = x$ be the CDF and $f(x) = 1$ be the PDF of a generic X_i.

Let's first focus on the maximum M and let $G(M)$ be its CDF and $g(M)$ be its PDF. Exploiting the fact that the random variables are i.i.d. we obtain that:

$$
\begin{aligned}
G(M) &= [F(M)]^n = M^n \\
g(M) &= n \times [F(M)]^{n-1} \times f(M) = n \times M^{n-1} \\
E[M] &= \int_0^1 M \times n \times M^{n-1} = n \int_0^1 M^n \\
&= n \times [\frac{M^{n+1}}{n+1}]_0^1 \\
&= \frac{n}{n+1}
\end{aligned}
$$

Let's now focus on the minimum m and let $H(m)$ be its CDF and $h(m)$ be its PDF. Exploiting the fact that the random variables are i.i.d. we obtain that:

$$
\begin{aligned}
1 - H(m) &= [1 - F(m)]^n \\
H(m) &= 1 - [1 - F(m)]^n = 1 - [1 - m]^n \\
h(m) &= n[1 - m]^{n-1} \\
E[m] &= \int_0^1 m \times n[1 - m]^{n-1} \\
&= \frac{1}{n+1}
\end{aligned}
$$

Exercise 9

Part a

We want to learn $P(|S_n - 500| \leq 20)$, since the 2% of 1000 equals 20, where $S_n = \sum_{i=1}^n X_i$ is the total number of good states, $X_i \sim Bern(p)$, with $p = 0.5$ and $n = 1000$. Wlog, we define $X_i = 1$ when the economy is in the good state.
Notice that the following events are equivalent:

$$
\begin{aligned}
|S_n - 500| &\leq 20 \\
480 \leq S_n &\leq 520 \\
\frac{480}{1000} \leq \frac{S_n}{1000} &\leq \frac{520}{1000} \\
0.48 - 0.5 \leq \frac{S_n}{1000} - 0.5 &\leq 0.52 - 0.5 \\
\frac{-0.02}{0.25} \leq Z_n &\leq \frac{0.02}{0.25}
\end{aligned}
$$

where Z_n is the standardized version of $\frac{S_n}{n} = \bar{X}_n$, i.e. the sample mean. We could compute directly the sampling distribution of Z_n. However, it is much more convenient to approximate it with a standard normal thanks to the CLT. Therefore we have that:

$$\begin{aligned} P(|S_n - 500| \leq 20) &= P(-0.08 \leq Z_n \leq 0.08) \\ &= \Phi(0.08) - \Phi(-0.08) = 0.0638 \end{aligned}$$

Part b

Let $\bar{X}_n = S_n/n$ from the previous part, which is the sample mean, but also the proportion of good states realized in the sample. Notice that $E[\bar{X}_n] = np/n = p = 0.5$, and $V[\bar{X}_n] = p(1-p)/n$. By Chebyshev's Inequality:

$$P(|\bar{X}_n - 0.5| \leq 0.01) > 1 - \frac{0.25}{0.01^2 \times n}$$

Therefore, we want to find n such that:

$$\frac{0.25}{0.0001 \times n} = 0.01$$

Therefore, the number of periods necessary is equal to $n = 250000$.

Exercise 10

Notice that $e(.)$ is a convex function. Therefore, by Jensen's Inequality:

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

Therefore it must be that $E[\theta X] = E[X]\theta < 0$. Since $E[X] < 0$, it must be that $\theta > 0$.