Math, Problem Set #1, Probability Theory

Elysa Strunin

June 28, 2017

1. Chapter 3 Exercises

3.6

$\forall i \in I$, $P(A \cap B_i) = P(A|B_i)P(B_i)$ by Bayes
$P(A) = \sum_{i \in I} P(A|B_i)P(B_i)$ by nature of partition and law of total probability
$\forall A \in F$, have $P(A) = \sum_{i \in I} P(A \cap B_i)$

3.8

$P(\bigcup_{k=1}^{n} E_k)^c = (\bigcap_{k=1}^{n} E_k^c)$ by deMorgan
$P(\bigcup_{k=1}^{n} E_k) = 1 - (\bigcap_{k=1}^{n} E_k^c)$ by complement
$= 1 - \prod_{k=1}^{n} P(E_k^c)$ by independence
$= 1 - \prod_{k=1}^{n}(1 - P(E_k))$

3.11

P(s = crime | s tested +) = P(s = crime , s tested +) / P(s tested +) by Bayes
= P(s tested + | s = crime) P(s = crime)/ P(s tested +) by Bayes = (1) (1/250 million) / (1/3 million)
= 0.012

3.12

P(win | stays) = 1/3
P(win | changes) = 1/2
Likewise, if 10 doors and 8 opened doors:
P(win | stays) = 1/10
P(win | changes) = 1/2

3.16

$var[X] = E[(X - \mu)^2]$ by definition
$= E[X^2 - 2X\mu + \mu^2]$
$= E[X^2] - 2E[X\mu] + E[\mu^2]$ by linearity of expectation
$= E[X^2] - 2\mu^2 + \mu^2$ by nature of μ
$= E[X^2] - \mu^2$

3.33

$\forall \epsilon > 0, P(|\frac{X_1 + X_2 + ... + X_n}{n} - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2}$ by WLLN
$B = Binomial(n, p) \Rightarrow B = \sum_{i=1}^{n} X_i$ where $X_i = Bernoulli(p)$

$\Rightarrow$ here $\mu = p$ and $\sigma^2 = p(1-p)$
$\forall \epsilon > 0, P(|\frac{B}{n} - p| \geq \epsilon) \leq \frac{p(1-p)}{n\epsilon^2}$

3.36

By CLT and Bernoulli distribution properties,
$P(\frac{5500-5000}{\sqrt{(0.801)(1-0.801)6242}} \leq y) \longrightarrow \Phi(y)$
So $P(event) = 1 - \Phi(15.85) \approx 0$

2

a

Events A, B, C that maintain pairwise independence but not mutual independence (classic example):
Let A and B be two independent coin tosses (1 for heads, 0 for tails)
Let $C = 1$ if exactly one of those tosses produced heads, and 0 otherwise
$P(A = 0) = P(B = 0) = P(C = 0) = P(A = 1) = P(B = 1) = P(C = 1) = 0.5$
$P(A \cap B) = P(A \cap C) = P(B \cap C) = 0.25$ for all intersections
$\Rightarrow$ pairwise independence is maintained throughout
But $P(A \cap B \cap C) = 0.25$ for the observed intersections, which $\neq P(A)P(B)P(C) = 0.125$
$\Rightarrow$ the events are not mutually independent (which makes sense because A and B determine C)

b

Let Ω be a set of 8 equally likely points.
A, B, C encompass 4 points each $\Rightarrow P(A) = P(B) = P(C) = 0.5$
To communicate the layout, can label the events in a 4x2 grid:
A = (2:3)x(1:2)
B = (3:4)x(1:2)
C = (1:3)x(2), (1,1)
$\Rightarrow P(A \cap B \cap C)$ = one point $= 0.125 = P(A)P(B)P(C)$
And $P(A \cap B) = P(A \cap C) = 0.25$
But $P(B \cap C) = 0.125 \neq P(B)P(C)$

3

Benford's Law is a well-defined discrete probability distribution:
$P(d$ being first digit $) = P(d) = log_{10}(1 + \frac{1}{d})$, $d \in 1, 2, 3, ..., 9$
Axioms (1) and (3) are fulfilled trivially
Axiom (2): $\sum_{d=1}^{9} log_{10}(1 + \frac{1}{d}) = 1$
$\sum_{d=1}^{9} log_{10}(1 + \frac{1}{d}) = log_{10}(\frac{2}{1} * \frac{3}{2} * \frac{4}{3} * ... * \frac{10}{9}) = log_{10}(\frac{10!}{9!}) = log_{10}(10) = 1$
$\Rightarrow$ Axiom 2 is fulfilled

4

a

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

b

$E[lnX] = \sum_{i=1}^{\infty} p(X_i)lnX_i$
$= \sum_{i=1}^{\infty} (\frac{1}{2^i})ln(2^i)$
$= ln2 \sum_{i=1}^{\infty} i(\frac{1}{2^i})$
(Not sure how to calculate this further)

5

Let $a + b = 1$, where a, b are the fractions of wealth invested in U.S. and Swiss currency, respectively
Let X, Y be investment amounts of U.S., Swiss investors, respectively
U.S. investor seeks to maximize $E[X]$:
$E[X] = 0.5(Xa + 1.25Xb) + 0.5(Xa + 0.8Xb)$
Substituting in the constraint, solve for a
$E[X] = 1.025 - 0.025a$
Decreasing in a, so U.S. investor should choose $a = 0$
$\Rightarrow$ U.S. investor invests only in Swiss francs
The problem is symmetric for the Swiss investor
$\Rightarrow$ Swiss investor invests only in USD

6

a

$E[X] < \infty$, $E[X^2] = \infty$:
Consider random variable X on the interval $[1, \infty)$ with pdf $f_x = \frac{2}{x^3}$
Continuous interval, so can map to Ω
Integrates to 1 over the interval
$E[X] = \int_1^{\infty} x\frac{2}{x^3}dx = 2 < \infty$
$E[X^2] = \int_1^{\infty} x^2\frac{2}{x^3}dx = 2ln(\infty) = \infty$

7

a

$Y = X$ with probability 0.5
$Y = -X$ with probability 0.5
X is symmetric about 0 and, more specifically, $\sim N(0, 1)$
$Y \sim N(0, 1)$

b

$\forall X_n, Y_n$:
$Y_n = X_n$ or $Y_n = -X_n$
$|Y| = |X|$
$\Rightarrow P(|X| = |Y|) = 1$

c

$Y = XZ$
Y is dependent on X
$\Rightarrow X, Y$ are not independent
Counterexample for independence: If independent, $P(Y = y) = P(Y = y|X = x)$
Take $P(Y = -1) = 0$ because continuous
But $P(Y = -1|X = -1) = 0.5 \Rightarrow X, Y$ are not independent

d

$cov(XY) = E[XY] - E[X]E[Y]$
$E[X] = E[Y] = 0$ from (a)
$cov(XY) = E[XY]$
Clear that XY remains symmetric around 0 $\Rightarrow E[XY] = 0 \Rightarrow cov(X, Y) = 0$

e

False. This exercise provides a counterexample:
Normally distributed with zero covariance, but not independent

8

m:
$P(m = x) = P($ all other $X_i > x) = C(1 - x)^{n-1}$, where pdf must integrate to 1
$\Rightarrow C = n \Rightarrow f_m = n(1 - x)^{n-1}$
cdf: $P(m \leq a) = -\frac{1}{n}(1 - a)^n + \frac{1}{n}$ (integrated pdf over the interval)
$E[m] = n \int_0^1 x(1 - x)^{n-1}$
M: Similarly,
$P(M = x) = P($ all other $X_i < x) = C(x)^{n-1}$, where pdf must integrate to 1
$\Rightarrow f_M = n(x)^{n-1}$
cdf: $P(M \leq a) = a^n$ (integrated pdf over the interval)
$E[M] = n \int_0^1 x^n = \frac{n}{n+1}$

9

Consider "good"=1, bad=0, state $\sim Bernoulli(p = 0.5)$ and $\sigma = \sqrt{p(1 - p)} = 0.5$

a

CLT: $P(\frac{490-500}{0.5*\sqrt{1000}} \leq y) = \Phi(y) = 0.264$
Since 2-tailed, probability of the specified extremes $= 2 * 0.264 = 0.527$
So desired probability $= 1 - 0.527 = 0.47$

b

Chebyshev and WLLN:
$P(|$ proportion of good states - $\mu| \geq 0.01) \leq \sigma^2/n\epsilon^2 = 1 - 0.99 = 0.01$
Solve for $n : 0.01 = 0.25/n(0.01)^2$
$n = 250,000$

10

$E[X] < 0 \Rightarrow \int xp(x)dx < 0$ where $\int p(x)dx = 1$
$E[e^{\theta x}] = 1 \Rightarrow \int e^{\theta x}p(x)dx = 1 \Rightarrow e^{\theta x}p(x)$ is a pdf
Can see that if $p(x) = -\theta$, this is the exponential distribution
For exponential here, $E[X] = -1/\theta$
$E[X] < 0 \Rightarrow -1/\theta < 0 \Rightarrow \theta > 0$