Section 6 Solutions

Based on the work of many CS109 instructors and course staff members.

1 Warmups

1.1 Sums of Random Variables

For each $X$ and $Y$ below, let $X$ and $Y$ be independent.

1. Let $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$. What is $\mu$ and $\sigma^2$ for $X + Y \sim \mathcal{N}(\mu, \sigma^2)$?

2. Let $X \sim \text{Uni}(0, 1)$ and $Y \sim \text{Uni}(0, 1)$. What is the PDF for $X + Y$?

3. In general, two random variables $X$ and $Y$, what is the PDF $f$ of $X + Y$?

Answer.

1. $\mu = \mu_1 + \mu_2$ and $\sigma^2 = \sigma_1^2 + \sigma_2^2$. How convenient!

2. $f_{X+Y}(a) = \begin{cases} a & 0 \leq a \leq 1 \\ 2 - a & 1 \leq a \leq 2 \\ 0 & \text{otherwise} \end{cases}$

3. $f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y) f_Y(y) dy$

   It is good to remember these equations, but perhaps another message from lecture that it is
difficult to sum random variables. The derivation for Uniform distributions is difficult. And
solving for the general random variables is even worse. But we can pick distributions, like the
Normal distribution, that are easy to use!

1.2 General Inference

Suppose $X_1, \ldots, X_4$ are discrete random variables. We will abuse notation and write $p(x_1, x_2, x_3, x_4)$
to represent $P(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4)$. In your answers, feel free to do the same. For
example, $p(x_1, x_3) = P(X_1 = x_1, X_3 = x_3)$. For the following cases, decompose into four terms, with
each being as simple as possible.

1. If there is no assumption of independence, what is $p(x_1, x_2, x_3, x_4)$?

2. If all variables are assumed independent, what is $p(x_1, x_2, x_3, x_4)$?

3. Assuming the variables follow the Bayesian network structure below, what is $p(x_1, x_2, x_3, x_4)$?
Answer.

1. \( p(x_1)p(x_2|x_1)p(x_3|x_1,x_2)p(x_4|x_1,x_2,x_3) \) (Chain Rule)
2. \( p(x_1)p(x_2)p(x_3)p(x_4) \)
3. \( p(x_1)p(x_2)p(x_3|x_1,x_2)p(x_4|x_3) \)

2 Problems

2.1 Fish Sticks (courtesy of Lisa Yan)

Fish Sticks, the online platform designed to meet all of your fish stick needs, wants to model their hourly homepage traffic from Stanford. The company decides to model two different behaviors for homepage visits according to the Bayesian Network on the right:

\[
\begin{align*}
D & \sim \text{Ber}(p = 0.3) \\
A & \sim \text{Poi}(5) \\
B & \sim \text{Poi}(3) \\
\end{align*}
\]

\[
\begin{align*}
D = 0: & \quad \text{Poi}(5) \\
D = 1: & \quad \text{Poi}(8) \\
\end{align*}
\]

A \sim \text{Poi}(5) \quad D \sim \text{Ber}(p = 0.3) \quad B \sim \text{Poi}(3) 

D = 0: \quad \text{Poi}(3) \\
D = 1: \quad \text{Poi}(6) \\

A \sim \text{Poi}(5) \\
D \sim \text{Ber}(p = 0.3) \\
B \sim \text{Poi}(3) 

A and B are the numbers of Stanford students and faculty, respectively, who visit the Fish Sticks homepage in an hour. Since Fish Sticks does not know when Stanford people eat, the company models demand as a "hidden" Bernoulli random variable \( D \), which determines the distribution of \( A \) and \( B \). Recall that in a Bayesian Network, random variables are conditionally independent given their parents. For example, given \( D = 0 \), \( A \sim \text{Poi}(5) \) and \( B \sim \text{Poi}(3) \), two independent random variables.

a. Given that 6 users from group \( A \) visit the homepage in the next hour, what is the probability that \( D = 0 \)?

b. What is the probability that in the next hour, the total number of users who visit the homepage from groups \( A \) and \( B \) is equal to 12, i.e., what is \( P(A + B = 12) \)?

c. Now simulate \( P(A + B = \text{total}) \), where \( \text{total} = 12 \), by implementing the \text{infer_prob_total}(\text{total}, \text{ntrials}) \) function below using rejection sampling.
**Answer.**

a. Note that given $D = 0$, $A \sim \text{Poi}(\lambda = 5)$, and given $D = 1$, $A \sim \text{Poi}(\lambda = 8)$. By Bayes’ Theorem,

$$P(D = 0|A = 6) = \frac{P(A = 6|D = 0)P(D = 0)}{P(A = 6|D = 0)P(D = 0) + P(A = 6|D = 1)P(D = 1)}$$

$$= \frac{\frac{5^6 e^{-5}}{6!} (1 - 0.3)}{\frac{5^6 e^{-5}}{6!} (1 - 0.3) + \frac{8^6 e^{-8}}{6!} (0.3)}$$

$$= \frac{5^6 e^{-5} (1 - 0.3)}{5^6 e^{-5} (1 - 0.3) + 8^6 e^{-8} (0.3)} \approx 0.7364$$

b. By Law of Total Probability,

$$P(A + B = 12) = P(A + B = 12|D = 0)P(D = 0) + P(A + B = 12|D = 1)P(D = 1).$$

$A$ and $B$ are conditionally independent Poisson random variables given $D$, and therefore $A + B|D = 0 \sim \text{Poi}(\lambda = 8)$ and $A + B|D = 1 \sim \text{Poi}(\lambda = 14)$. Using the Poisson PMF,

$$P(A + B = 12) = \frac{8^{12} e^{-8}}{12!} \cdot (1 - 0.3) + \frac{14^{12} e^{-14}}{12!} \cdot (0.3) \approx 0.0632.$$  

c. Here’s one idea:

```python
import numpy as np
from scipy import stats

def infer_prob_total(total, ntrials):
    n_samples_event = 0
    for i in range(ntrials):
        d = stats.bernoulli.rvs(0.3)
        user_sum = 0
        if d == 0:
            user_sum += stats.poisson.rvs(5) + stats.poisson.rvs(3)
        else:
```
user_sum += stats.poisson.rvs(8) + stats.poisson.rvs(6)

if user_sum == 12:
    n_samples_event += 1

prob = n_samples_event/ntrials
return prob

ntrials = 50000
total = 12
print("Simulated P(A + B) = ", infer_prob_total(total, ntrials))