Quiz Section 7

Review

1) Central Limit Theorem (CLT): Let $X_1, \ldots, X_n$ be iid random variables with $\mathbb{E}[X_i] = \mu$ and $\text{Var}(X_i) = \sigma^2$. Let $X = \sum_{i=1}^{n} X_i$, which has $\mathbb{E}[X] = n\mu$ and $\text{Var}(X) = n\sigma^2$. Let $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, which has $\mathbb{E}[\overline{X}] = \mu$ and $Var(\overline{X}) = \frac{\sigma^2}{n}$. $\overline{X}$ is called the *sample mean*. Then, as $n \to \infty$, $\overline{X}$ approaches the normal distribution $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$. Standardizing, this is equivalent to $Y = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}$ approaching $\mathcal{N}(0,1)$. Similarly, as $n \to \infty$, X approaches $\mathcal{N}(n\mu, n\sigma^2)$ and $Y' = \frac{X-n\mu}{\sigma\sqrt{n}}$ approaches $\mathcal{N}(0,1)$.

It is no surprise that $\overline{X}$ has mean μ and variance σ^2/n – this can be done with simple calculations. The importance of the CLT is that, for large n, regardless of what distribution X_i comes from, $\overline{X}$ is *approximately normally distributed with mean μ and variance σ^2/n*. Don't forget the continuity correction, only when $X_1, \ldots, X_n$ are discrete random variables.

2) Multivariate: Discrete to Continuous:

	Discrete	**Continuous**
Joint PMF/PDF	$p_{X,Y}(x,y) = \mathbb{P}(X = x, Y = y)$	$f_{X,Y}(x,y) \neq \mathbb{P}(X = x, Y = y)$
Joint range/support $\Omega_{X,Y}$	$\{(x,y) \in \Omega_X \times \Omega_Y : p_{X,Y}(x,y) > 0\}$	$\{(x,y) \in \Omega_X \times \Omega_Y : f_{X,Y}(x,y) > 0\}$
Joint CDF	$F_{X,Y}(x,y) = \sum_{t \leqslant x, s \leqslant y} p_{X,Y}(t,s)$	$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(t,s)\, ds dt$
Normalization	$\sum_{x,y} p_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx dy = 1$
Marginal PMF/PDF	$p_X(x) = \sum_y p_{X,Y}(x,y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy$
Expectation	$\mathbb{E}[g(X,Y)] = \sum_{x,y} g(x,y) p_{X,Y}(x,y)$	$\mathbb{E}[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y) dx dy$
Independence must have	$\forall x, y, p_{X,Y}(x,y) = p_X(x) p_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$	$\forall x, y, f_{X,Y}(x,y) = f_X(x) f_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$

Task 1 – Round-off error

Let X be the sum of 100 real numbers, and let Y be the same sum, but with each number rounded to the nearest integer before summing. If the roundoff errors are independent and uniformly distributed between -0.5 and 0.5, what is the approximate probability that $|X - Y| > 3$?

Task 2 – Tweets

A prolific Twitter user tweets approximately 350 tweets per week. Let's assume for simplicity that the tweets are independent, and each consists of a uniformly random number of characters between 10 and 140. (Note that this is a discrete uniform distribution.) Thus, the central limit theorem (CLT) implies that the number of characters tweeted by this user is approximately normal with an appropriate mean and variance. Assuming this normal approximation is correct, estimate the probability that this user tweets between 26,000 and 27,000 characters in a particular week. (This is a case where continuity correction will make virtually no difference in the answer, but you should still use it to get into the practice!).

Task 3 – Joint PMF's

Suppose X and Y have the following joint PMF:

X/Y	1	2	3
0	0	0.2	0.1
1	0.3	0	0.4

a) Identify the range of X (Ω_X), the range of Y (Ω_Y), and their joint range ($\Omega_{X,Y}$).

b) Find the marginal PMF for X, $p_X(x)$ for $x \in \Omega_X$.

c) Find the marginal PMF for Y, $p_Y(y)$ for $y \in \Omega_Y$.

d) Are X and Y independent? Why or why not?

e) Find $\mathbb{E}[X^3 Y]$.

Task 4 – Do You "Urn" to Learn More About Probability?

Suppose that 3 balls are chosen without replacement from an urn consisting of 5 white and 8 red balls. Let $X_i = 1$ if the i-th ball selected is white and let it be equal to 0 otherwise. Give the joint probability mass function of

a) X_1, X_2

b) X_1, X_2, X_3

Task 5 – Continuous joint density

The joint density of X and Y is given by

$$f_{X,Y}(x,y) = \begin{cases} xe^{-(x+y)} & x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

and the joint density of W and V is given by

$$f_{W,V}(w,v) = \begin{cases} 2 & 0 < w < v, 0 < v < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Are X and Y independent? Are W and V independent?

Task 6 – Trapped Miner

A miner is trapped in a mine containing 3 doors.

- D_1: The 1^{st} door leads to a tunnel that will take him to safety after 3 hours.
- D_2: The 2^{nd} door leads to a tunnel that returns him to the mine after 5 hours.
- D_3: The 3^{rd} door leads to a tunnel that returns him to the mine after a number of hours that is Binomial with parameters $(12, \frac{1}{3})$.

At all times, he is equally likely to choose any one of the doors. What is the expected number of hours for this miner to reach safety?

Task 7 – Lemonade Stand

Suppose I run a lemonade stand, which costs me \$100 a day to operate. I sell a drink of lemonade for \$20. Every person who walks by my stand either buys a drink or doesn't (no one buys more than one). If it is raining, n_1 people walk by my stand, and each buys a drink independently with probability p_1. If it isn't raining, n_2 people walk by my stand, and each buys a drink independently with probability p_2. It rains each day with probability p_3, independently of every other day. Let X be my profit over the next week. In terms of n_1, n_2, p_1, p_2 and p_3, what is $\mathbb{E}[X]$?