Stats 117 Problem Set 4

Due: Monday, April 27 5:00 p.m. on Gradescope

Please show your work for each exercise. If you collaborate with someone else—this is fine—be sure to note that in your homework submission. You must each write up separate answer sets. Any starred exercise is optional: they are extra challenging theoretical exercises for further developing your mastery.

Question 4.1: Let $X \sim \mathsf{Geom}(p)$ and $t \in \mathbb{R}$ be fixed.

(a) Compute $M(t) = \mathbb{E}[e^{tX}]$ as a function of t.

(b) Use this to compute $\mathbb{E}[Xe^{tX}]$ as a function of t. *Hint.* What is $\frac{\partial}{\partial t} e^{ta}$?

Question 4.2: In bridge, each of four players has 13 of the 52 cards in a standard card deck; pairs of players (identified as North/South and East/West) are teammates (we call this a "partnership"), and there are 13 cards of each suit (spades, hearts, diamonds, clubs). We say a partnership is *void in a suit* if between its two players there are no cards in a particular suit. The deck is shuffled between hands, so the probability the opposing partnership is void is independent from hand to hand.

(a) What is the probability that our opponents have a void in any suit (it does not matter which)?

(b) We open our hand, and see that we have 4 spades. What is the probability that the opponents are void in spades?

(c) We open our hand, and we observe that have we have 8 spades. What is the probability that the opponents are void in spades?

(d) We play 1000 hands of bridge. What is the probability that the opposing partnership has a void of a suit in at least one of those 1000 hands?

Question 4.3: Let $X \sim \mathsf{Poisson}(\lambda)$.

(a) Compute the function $M(t) := \mathbb{E}[e^{tX}]$ for $t \in \mathbb{R}$.

(b) Give $M'(0), M''(0)$, and $M'''(0)$.

Question 4.4: In most models of the utility of wealth for humans, utility is a *concave* function of the amount of money, meaning that there are diminishing returns: going from \$1 to \$2 makes one happier than going from \$1,000,000 to \$1,000,001. Consider now the St. Petersburg game, and let X be the wealth obtained by playing the game. (So we have that $X = 2^k$ with probability $1/2^{k+1}$ for $k = 0, 1, 2, \ldots$.) We consider a few different utility functions to define our "happiness" at achieving wealth X.

(a) Take the $\sqrt{\cdot}$ function and let $H = \sqrt{X}$. Compute $\mathbb{E}[H]$.

(b) Take the $\log_2$ function and let $H = \log_2 X$. Compute $\mathbb{E}[H]$.

Question 4.5: It is natural to model counts of organisms in an area via the Poisson distribution (we will see why this is the case later in class). Suppose that we collect a volume V of seawater. The number of non-single-celled organisms in this volume follows a $\mathsf{Poisson}(\mu)$ distribution, where $\mu = 10V$.

(a) What is the probability there are at least 12 organisms in $V = 1.5$ cubic meters of water?

(b) How large a volume V must we collect to have a probability of at least .999 of finding at least one organism?

Question 4.6* (The Poisson approximation)**:** If $X \sim \mathsf{Binomial}(n,p)$ and np is "not too large," then by taking $\lambda = np$, that is, for $p = \frac{\lambda}{n}$, the random variable $Y \sim \mathsf{Poisson}(\lambda)$ provides a good approximation to X in that

$$P(X = i) \approx P(Y = i) \tag{4.1}$$

when i is "not too large." In this question, we will make this approximation rigorous in the sense that if $np \ll \sqrt{n}$ and $i \ll \sqrt{n}$, the approximation (4.1) is accurate to within a small multiplicative factor.

(a) Show that if $f(x) = \log(1 - x)$ and $g(x) = -x - x^2$, then $f(x) \ge g(x)$ for all $x \in [0, \frac{1}{4}]$. Conclude that

$$\exp(-x - x^2) \le 1 - x \le e^{-x},$$

where the first equality is valid when $0 \le x \le \frac{1}{4}$, while the second is valid for all x (you do not need to prove the second inequality).

(b) Show that if $\lambda = pn$, i.e., $p = \frac{\lambda}{n}$ then

$$P(X = i) = \frac{n(n-1)\cdots(n-i+1)}{n^i}\left(1 - \frac{\lambda}{n}\right)^{n-i} \cdot \frac{\lambda^i}{i!}.$$

(c) Assume that for some $\epsilon > 0$, we have $\frac{\lambda i}{n} \le \epsilon$. Show that

$$P(X = i) \le \frac{e^{-\lambda}\lambda^i}{i!} \cdot e^{\epsilon}.$$

(d) Assume that for some $0 < \epsilon \le \frac{1}{4}$, we have $\frac{\lambda}{n} \le \epsilon$, $\frac{i^2}{n} \le \epsilon$, and $\frac{\lambda^2}{n} \le \epsilon$. Show that

$$P(X = i) \ge \frac{e^{-\lambda}\lambda^i}{i!} \cdot e^{-2\epsilon - \epsilon^2}.$$

(e) Combining the preceding results, conclude that for any sequences of random variables $X_{n,p} \sim \mathsf{Binomial}(n,p)$ and $Y_{n,p} \sim \mathsf{Poisson}(n \cdot p)$ and any sequence ϵ_n satisfying $\epsilon_n \to 0$ as $n \to \infty$,

$$\max_i \max_p \left\{ \left| \log \frac{P(X_{n,p} = i)}{P(Y_{n,p} = i)} \right| \text{ such that } 0 \le p \le \frac{\epsilon_n}{\sqrt{n}},\ 0 \le i \le \epsilon_n \sqrt{n} \right\} \to 0.$$

Conclude that (roughly) as long as p smaller than $1/\sqrt{n}$ and i is smaller than $\sqrt{n}$, the approximation (4.1) is valid for $X \sim \mathsf{Binomial}(n,p)$ and $Y \sim \mathsf{Poisson}(np)$.