STATS 116: Practice midterm exam

Monday, July 24, 2023 from 10:30 am to 11:20 am PDT

Name ____

SUNET ID ____

There are 4 problems on this examination, worth a total of 48 points. The point value of each problem is given below. Subparts within a problem may not carry equal weight. You are not expected to completely solve all of the problems within the time limit, so do your best.

This examination is closed book, with the exception of one standard size (8.5 inch by 11 inch) sheet of paper which may contain any information placed on it prior to the start of the examination. As per the Honor Code and syllabus, any collaboration with other students or any other individuals is strictly prohibited, as is the use of electronic devices during the examination (other than to check the time).

If a problem subpart depends on the answer to a previous subpart, you may receive full credit for this subpart without solving the previous subpart by expressing your answer in terms of the answer to the previous subpart. You may use the back page of each problem if you need more space; please indicate this on the main page with the problem if you do so, to reduce the probability that it is missed during grading. Unless you are explicitly asked to simplify, you may leave your answer in terms of binomial coefficients. Even if you are asked to simplify, you will receive most of the credit with a correct answer that you are unable to simplify either algebraically or via a "story." Good luck!

Problem 1	__________	out of 12
Problem 2	__________	out of 12
Problem 3	__________	out of 12
Problem 4	__________	out of 12
Total	__________	out of 48

Problem 1

A certain baby cries if and only if she is hungry, tired, or both. Let C be the event that the baby is crying, H be the event that she is hungry, and T be the event that she is tired. Let $P(C) = c$, $P(H) = h$, and $P(T) = t$, where none of c, h, t are equal to 0 or 1. Suppose that H and T are independent.

(a) Find c, in terms of h and t. (Simplify fully.)

(b) Find $P(H \mid C)$, $P(T \mid C)$, and $P(H, T \mid C)$ (in terms of c, h, t; simplify fully).

(c) Are H and T conditionally independent given C? Justify your answer in two ways: algebraically using the quantities from (b), and with a brief but clear intuitive explanation in words.

Problem 2

A publishing company employs two proofreaders, Prue and Frida. When Prue is proofreading a book, for each typo she has probability p of catching it and $q = 1 - p$ of missing it, independently. When Frida is proofreading a book, for each typo she has probability f of catching it and $g = 1 - f$ of missing it, independently.

(a) A certain book draft has n typos. The company randomly assigns it to one of the two proofreaders, with equal probabilities. Specify the distribution of the number of typos that get detected.

(b) Another book is being written. When a draft of the book is complete, it will have a Pois(λ) number of typos, and will be assigned to Prue to proofread. Find the probability that Prue catches exactly k typos, given that she misses exactly m typos (assume k and m are nonnegative integers).

Problem 3

The United States Senate has 100 senators: 2 from each of the 50 states. Each state has a junior senator and a senior senator (based on which of them has served longer). A committee of size 20 is formed randomly, with all sets of 20 senators equally likely.

(a) Specify the distribution of the number of junior senators on the committee.

(b) Find the expected number of junior senators on the committee (simplify).

(c) Find the expected number of states such that both senators from that state are on the committee (simplify).

Problem 4

Let $N \sim \text{FS}(p)$, with $\frac{1}{2} < p < 1$.

(a) Find $\mathbb{E}[2^N]$ (simplify).

(b) Explain in words why the answer to (a) makes sense in the extreme cases where p is very close to $1/2$ and where p is very close to 1. (You can do this part even if you haven't solved (a), by giving an intuitive argument for what you think the answer should look like in these extreme cases.)

(c) Let F be the CDF of the indicator random variable for the event $N > 3$. Find F, being sure to specify $F(x)$ for all real x.

Table of distributions

Below is some information about some named distribution families that might be useful. Note: below, the letter q denotes the quantity $1-p$.

Name	Param.	PMF	Mean	Variance
Bernoulli	p	$P(X=1)=p, P(X=0)=q$	p	pq
Binomial	n, p	$\binom{n}{k} p^k q^{n-k}$, for $k \in \{0, 1, \ldots, n\}$	np	npq
FS	p	pq^{k-1}, for $k \in \{1, 2, \ldots\}$	$1/p$	q/p^2
Geom	p	pq^k, for $k \in \{0, 1, 2, \ldots\}$	q/p	q/p^2
NBinom	r, p	$\binom{r+n-1}{r-1} p^r q^n, n \in \{0, 1, 2, \ldots\}$	rq/p	rq/p^2
HGeom	w, b, n	$\frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}}$, for $k \in \{0, 1, \ldots, n\}$	$\mu = \frac{nw}{w+b}$	$(\frac{w+b-n}{w+b-1})\mu(1-\frac{\mu}{n})$
Poisson	λ	$\frac{e^{-\lambda}\lambda^k}{k!}$, for $k \in \{0, 1, 2, \ldots\}$	λ	λ