Probability & Statistics
Fall 2021
Final Exam
Jan. 13, 2022
Time Limit: 150 Minutes

Name (Print): ____________________

Advisor Name ____________________

This exam contains 12 pages (including this cover page) and 9 problems. Check to see if any pages are missing. Enter all requested information on the top of this page, and put your initials on the top of every page, in case the pages become separated.

Try to answer as many problems as you can. The following rules apply:

- **Mysterious or unsupported answers will not receive full credit**. A correct answer, unsupported by calculations, explanation, or algebraic work will receive no credit; an incorrect answer supported by substantially correct calculations and explanations might still receive partial credit.

Do not write in the table to the right.

Problem	Points	Score
1	10	
2	10	
3	10	
4	15	
5	10	
6	15	
7	10	
8	10	
9	10	
Total:	100	

1. (10 points) Basic Concept.
   (a) (5 points) Please describe the difference and connection between probability and statistics.
   (b) (5 points) Please describe the pros and cons of Bayesian statistical inference and classical statistical inference.

2. (10 points) Let X be a discrete r.v. whose distinct possible values are $x_0, x_1, \ldots,$ and let $p_k = P(X = x_k)$. The entropy of X is $H(X) = \sum_{k=0}^{\infty} p_k \log_2(1/p_k)$.

   (a) (5 points) Find $H(X)$ for $X \sim Geom(p)$.

   (b) (5 points) Let X and Y be i.i.d. discrete r.v.s. Show that $P(X = Y) \geq 2^{-H(X)}$.

3. (10 points) Let $X_1 \sim Expo(\lambda_1)$, $X_2 \sim Expo(\lambda_2)$ and $X_3 \sim Expo(\lambda_3)$ be independent.

   (a) (5 points) Find $E(X_1 + X_2 + X_3 | X_1 > 1, X_2 > 2, X_3 > 3)$ in terms of $\lambda_1, \lambda_2, \lambda_3$.

   (b) (5 points) Find $P(X_1 = min(X_1, X_2, X_3))$.

4. (15 points) Let $X \sim Gamma(a, \lambda)$, $Y \sim Gamma(b, \lambda)$. Assume X and Y are independent.
   (a) (5 points) Find the joint distribution of $T = X + Y$ and $W = \frac{X}{X+Y}$.
   (b) (5 points) Find the distribution of T and W respectively.
   (c) (5 points) Find $E(W)$.

5. (10 points) Instead of predicting a single value for the parameter, we give an interval that is likely to contain the parameter: A $1-\delta$ confidence interval for a parameter p is an interval $[\hat{p}-\epsilon, \hat{p}+\epsilon]$ such that $Pr\left(p \in [\hat{p}-\epsilon, \hat{p}+\epsilon]\right) \geq 1-\delta$. Now we toss a coin with probability p landing heads and probability $1-p$ landing tails. The parameter p is unknown and we need to estimate its value from experiment results. We toss such coin N times. Let $X_i = 1$ if the ith result is head, otherwise 0. We estimate p by using $\hat{p} = \frac{X_1+\ldots+X_N}{N}$. Find the $1-\delta$ confidence interval for p, then discuss the impacts of δ and N. **Hint**: You can use the following Hoeffding bound: Let the random variables $X_1, X_2, \ldots, X_n$ be independent with $E(X_i) = \mu$, $a \leq X_i \leq b$ for each $i = 1, \ldots, n$, where a, b are constants. Then for any $\epsilon \geq 0$,

$$\mathbb{P}(|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| \geq \epsilon) \leq 2e^{-\frac{2n\epsilon^2}{(b-a)^2}}.$$

6. (15 points) Given a coin with the probability p of landing heads. p is unknown and we need to estimate its value through data. In our data collection model, we have n independent tosses, result of each toss is either Head or Tail. Let X denote the number of heads in the total n tosses. Now we conduct experiments to collect data and find $X = k$. Then we need to find $\hat{p}$, the estimation of p.

   (a) (5 points) Assume p is an unknown constant. Find $\hat{p}$ through the MLE (Maximum Likelihood Estimation) rule.

   (b) (5 points) Assume p is a random variable with a prior distribution $p \sim Beta(a, b)$, where a and b are known constants. Find $\hat{p}$ through the MAP (Maximum a Posterior Probability) rule.

   (c) (5 points) Assume p is a random variable with a prior distribution $p \sim Beta(a, b)$, where a and b are known constants. Find $\hat{p}$ through the MMSE (Minimal Mean Squared Error) rule.

7. (10 points) We know that the MMSE of Y given X is given by $g(X) = E[Y|X]$. We also know that the Linear Least Square Estimate (LLSE) of Y given X, denoted by $L[Y|X]$, is shown as follows:

$$L[Y|X] = E(Y) + \frac{Cov(X,Y)}{Var(X)}(X - E(X)).$$

Now we wish to estimate the probability of landing heads, denoted by θ, of a biased coin. We model θ as the value of a random variable Θ with a known prior PDF $f_\Theta \sim \text{Unif}(0,1)$. We consider n independent tosses and let X be the number of heads observed.

(a) (5 points) Show that $E[(\Theta - E[\Theta|X])h(X)] = 0$ for any real function $h(\cdot)$.

(b) (5 points) Find the MMSE $E[\Theta|X]$ and the LLSE $L[\Theta|X]$. (Eve's law: $Var(Y) = E(Var(Y|X)) + Var(E(Y|X))$.)

8. (10 points) Show the following inequalities.

   (a) (5 points) Let $X \sim Pois(\lambda)$. If there exists a constant $a > \lambda$, then

   $$\mathbb{P}(X \geq a) \leq \frac{e^{-\lambda}(e\lambda)^a}{a^a}$$

   (b) (5 points) Let X be a random variable with finite variance σ^2. Then for any constant $a > 0$,

   $$\mathbb{P}(|X - \mathbb{E}[X]| \geq a) \leq \frac{2\sigma^2}{\sigma^2 + a^2}.$$

9. (10 points) Let $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$; X and Y are independent. Now let $Z_1 = sin(X + Y)$, $Z_2 = cos(X+Y)$.

   (a) (5 points) Find $E(Z_1)$ and $E(Z_2)$

   (b) (5 points) Find $Var(Z_1)$ and $Var(Z_2)$

Appendix: Bayes' Rule & LOTP

	Y **discrete**	Y **continuous**
X **discrete**	$P(Y=y\|X=x) = \frac{P(X=x\|Y=y)P(Y=y)}{P(X=x)}$	$f_Y(y\|X=x) = \frac{P(X=x\|Y=y)f_Y(y)}{P(X=x)}$
X **continuous**	$P(Y=y\|X=x) = \frac{f_X(x\|Y=y)P(Y=y)}{f_X(x)}$	$f_{Y\|X}(y\|x) = \frac{f_{X\|Y}(x\|y)f_Y(y)}{f_X(x)}$

	Y **discrete**	Y **continuous**
X **discrete**	$P(X=x) = \sum_y P(X=x\|Y=y)P(Y=y)$	$P(X=x) = \int_{-\infty}^{\infty} P(X=x\|Y=y)f_Y(y)dy$
X **continuous**	$f_X(x) = \sum_y f_X(x\|Y=y)P(Y=y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X\|Y}(x\|y)f_Y(y)dy$

Table of distributions

Name	**Param.**	**PMF or PDF**	**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 $k \in \{0,1,\dots,n\}$	np	npq
FS	p	$pq^{k-1},$ for $k \in \{1,2,\dots\}$	$1/p$	q/p^2
Geom	p	$pq^k,$ for $k \in \{0,1,2,\dots\}$	q/p	q/p^2
NBinom	r, p	$\binom{r+n-1}{r-1}p^r q^n, n \in \{0,1,2,\dots\}$	rq/p	rq/p^2
HGeom	w, b, n	$\frac{\binom{w}{k}\binom{b}{n-k}}{\binom{w+b}{n}},$ for $k \in \{0,1,\dots,n\}$	$\mu = \frac{nw}{w+b}$	$\left(\frac{w+b-n}{w+b-1}\right)n\frac{\mu}{n}\left(1-\frac{\mu}{n}\right)$
Poisson	λ	$\frac{e^{-\lambda}\lambda^k}{k!},$ for $k \in \{0,1,2,\dots\}$	λ	λ
Uniform	$a < b$	$\frac{1}{b-a},$ for $x \in (a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal	μ, σ^2	$\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/(2\sigma^2)}$	μ	σ^2
Log-Normal	μ, σ^2	$\frac{1}{x\sigma\sqrt{2\pi}}e^{-(\log x-\mu)^2/(2\sigma^2)}, x > 0$	$\theta = e^{\mu+\sigma^2/2}$	$\theta^2(e^{\sigma^2}-1)$
Expo	λ	$\lambda e^{-\lambda x},$ for $x > 0$	$1/\lambda$	$1/\lambda^2$
Gamma	a, λ	$\Gamma(a)^{-1}(\lambda x)^a e^{-\lambda x}x^{-1},$ for $x > 0$	a/λ	a/λ^2
Beta	a, b	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1},$ for $0 < x < 1$	$\mu = \frac{a}{a+b}$	$\frac{\mu(1-\mu)}{a+b+1}$
Chi-Square	n	$\frac{1}{2^{n/2}\Gamma(n/2)}x^{n/2-1}e^{-x/2},$ for $x > 0$	n	$2n$
Student-t	n	$\frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}(1+x^2/n)^{-(n+1)/2}$	0 if $n > 1$	$\frac{n}{n-2}$ if $n > 2$