Practice Final Exam

Math 170S (Winter 2025)
Instructor: Koffi Enakoutsa
TA: Colin Ni

Name: ____

Student ID: ____

Show all work. You may use one sheet (front and back) of notes. You may not use a calculator. Please ensure your phone is silenced and stowed away.

Duration: 3 hours.

The following is my own work, without the aid of any other person.

Signature: ____

Each problem is worth 25 points, for a total of 200 points, but there are 20 points of available extra credit.

Problem	**Points**	**Score**
1	25	
2	25	
3	25	
4	25	
5	25	
6	25	
7	25	
8	25	
Total	200	

Problem 1

Consider the following 20 sorted samples $X_1, \ldots, X_{20}$ drawn from the random variable $X \sim \text{Norm}(2, 1)$:

-0.1	0.13	0.6	0.77	1.28	1.39	1.51	1.57	1.65	1.74
2.04	2.04	2.06	2.15	2.53	2.6	2.73	3.05	3.23	3.69.

The sum of the samples and the sum of the squares of the samples are respectively

$$\sum_{i=1}^{20} X_i = 36.66 \qquad \text{and} \qquad \sum_{i=1}^{20} X_i^2 = 86.20.$$

(*i*) (10 points) Show that

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2.$$

(*ii*) (5 points) Use part (*i*) to compute the sample mean $\bar{X}$ and the sample variance s_X^2.

(*iii*) (10 points) Compute the third quartile $\pi_{0.75}$.

(*iv*) (5 points, extra credit) More generally, let $X_1, \ldots, X_n$ be samples drawn from a random variable X with variance $\text{Var}(X) = \sigma^2$. The sample variance s_X^2 is an estimator for σ^2, but so is

$$v_X := \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

In one or two sentences, explain the key difference between these two estimators and how it relates to Bessel's correction.

Problem 2

In this problem, we will explore the order statistics of the so-called power distribution. The *power distribution* $P(\alpha)$ *with power* $\alpha > 0$ is defined by the pdf

$$f_{P(\alpha)}(x) = \alpha x^{\alpha-1} \qquad \text{for } 0 \leq x \leq 1.$$

Let $P_1, \ldots, P_n$ be samples from $P(\alpha)$.

(i) (5 points) Check that $f_{P(\alpha)}$ is a valid pdf, *i.e.* that it integrates to 1.

(ii) (10 points) Determine the distribution of the nth order statistic $P_{(n)}$ in terms of a power distribution.

(iii) (10 points) The *Kumaraswamy distribution* $K(a, b)$ *with shape* $a, b \geq 0$ is defined by the pdf

$$f_{K(a,b)}(x) = abx^{a-1}(1 - x^a)^{b-1} \qquad \text{for } 0 \leq x \leq 1.$$

(Note the similarity with $\mathrm{Beta}(\alpha, \beta)$ which has pdf proportional to $x^{\alpha-1}(1 - x)^{\beta-1}$.) Determine the distribution of the 1st order statistic $P_{(1)}$ in terms of a Kumaraswamy distribution.

Problem 3

In this problem, we will find the MLE for the probability of success for a geometric random variable. Recall that a geometric distribution Geom(p) with probability of success $0 \leq p \leq 1$ has pdf

$$f_{\text{Geom}(p)}(x) = (1-p)^{x-1}p \qquad \text{for } x \geq 0.$$

(*i*) (5 points) Describe precisely the likelihood function.

(*ii*) (15 points) Find the maximum likelihood estimator $\hat{p}$ of the parameter p.

(*iii*) (5 points) Explain why the estimator $\hat{p}$ you found makes sense intuitively in terms of what a geometric random variable models and what p represents.

Problem 4

In this problem, we will find the method of moments estimators for the parameters of a Gamma distribution. Recall that a Gamma distribution $\mathrm{Gamma}(\alpha, \theta)$ has parameters $\alpha > 0$ and $\theta > 0$ and is defined by the pdf

$$f_{\mathrm{Gamma}(\alpha,\theta)}(x) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}} \qquad \text{for } x \geq 0.$$

Recall further that it has mean $\alpha\theta$ and variance $\alpha\theta^2$.

(i) (10 points) Let $X \sim \mathrm{Gamma}(\alpha, \theta)$. Find the method of moments estimators $\tilde{\alpha}$ and $\tilde{\theta}$ for the parameters α and θ respectively.

(ii) (5 points) Set

$$v_X^2 = \frac{n-1}{n} s_X^2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2 - \bar{X}^2.$$

Write $\tilde{\alpha}$ and $\tilde{\theta}$ in terms of V and $\bar{X}$.

(ii) (10 points) Method of moments estimators can sometimes produce nonsensical estimates. Provide a concrete example of this using the estimators $\tilde{\alpha}$ and $\tilde{\theta}$.

Problem 5

In this problem, we will work out an example of linear regression and discuss what happens when we introduce regularization. Consider the following dataset of x-values $x_1, \dots, x_8$ and y-values $y_1, \dots, y_8$:

x:	-2	4	-1	-2	0	3	1	1
y:	4	-7	3	5	2	-4	-1	-2.

(*i*) (5 points) Plot this dataset.

(*ii*) (20 points) Find the exact line of best fit using linear regression. (*Hint*: Your answer should contain the number 32/17.)

(*iii*) (5 points, extra credit) Recall that for the linear model $y(x) = \alpha + x\beta$, linear regression finds the parameters α and β that minimizes the mean squared error. Consider the regularized mean squared error function

$$\frac{1}{n}\sum_{i=1}^{n}(y(x_i) - y_i)^2 + \frac{\lambda}{n}(\alpha^2 + \beta^2)$$

where $\lambda > 0$ is some hyperparameter. This penalizes α and β being large, resulting in a simpler model. Find α and β that minimize this regularized error function. (*Hint*: Write the function as $\frac{1}{n}(\mathbf{X}\gamma - \mathbf{y})^T(\mathbf{X}\gamma - \mathbf{y}) + \frac{\lambda}{n}\gamma^T\gamma$ which has derivative $\frac{2}{n}(\mathbf{X}^T\mathbf{X}\gamma - \mathbf{X}^T\mathbf{y}) + \frac{2\lambda}{n}\gamma$.)

Problem 6

In this problem, we will give an example of a sufficient statistic and an example of a non-sufficient statistic. Let $X \sim \text{Pois}(\lambda)$ be an Poisson random variable with unknown rate $\lambda > 0$. Recall that the pdf is given by

$$f_{\text{Pois}(\lambda)}(k) = \frac{\lambda^k e^{-\lambda}}{k!} \qquad \text{for } k \geq 0 \text{ an integer}$$

and that the MLE of the parameter λ is the sample mean: $\hat{\lambda} = \bar{X}$.

(*i*) (25 points) Show that the MLE $\hat{\lambda}$ is a sufficient statistic for λ.

(*ii*) (5 points, extra credit) Show that the sample minimum is not a sufficient statistic for λ.

Problem 7

In this problem, we will walk you through how to determine the Bayes estimate of the mean of a normal distribution with a normal prior, and then we will discuss how to interpret the Bayes estimate. Let $X_1, \ldots, X_n$ be a sample from the normal random variable $\mathrm{Norm}(\mu, \sigma^2)$ where σ^2 is known, and assume that the prior distribution of μ is the normal distribution $\mathrm{Norm}(\mu_0, \sigma_0^2)$. We will be interested in the posterior distribution of μ given the sample mean $\bar{X}$.

(i) (5 points) Recall that the pdf of $\mathrm{Norm}(\mu, \sigma^2)$ is

$$f_{\mathrm{Norm}(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

Find a, b, c such that the joint distribution of $\bar{X}$ and μ is proportional to $\exp\left(\frac{a\mu^2 - b\mu}{c}\right)$.

(ii) (5 points) Observe that

$$\exp\left(-\frac{a\mu^2 - b\mu}{c}\right) = \exp\left(-\frac{(\mu - \frac{b}{2a})^2 - \frac{b^2}{4a^2}}{\frac{c}{a}}\right) \propto \exp\left(-\frac{(\mu - \frac{b}{2a})^2}{\frac{c}{a}}\right).$$

Find the normal distribution whose pdf this is proportional to.

(iii) (5 points) Deduce that the posterior distribution of μ given the sample mean $\bar{X}$ is

$$\mathrm{Norm}\left(\frac{\bar{X}\sigma_0^2 + \mu_0\sigma^2/n}{\sigma_0^2 + \sigma^2/n}, \frac{\sigma^2\sigma_0^2/n}{\sigma_0^2 + \sigma^2/n}\right).$$

(iv) (5 points) Let us take our Bayes estimate to be the mean of the posterior. Use this Bayes estimate to illustrate the idea that Bayes estimators are influenced by both the data and the prior.

(v) (5 points) Explain what it means for n to be large, and explain what happens to the Bayes estimate in this case. Similarly, explain what it means for σ_0^2 to be small, and explain what happens to the Bayes estimate in this case.

Problem 8

In this problem, we will construct a confidence interval for the difference between two means assuming unknown but proportional variances. Assume $X \sim \text{Norm}(\mu_X, \sigma^2)$ and $Y \sim \text{Norm}(\mu_Y, \sigma^2)$ are normal, where the means are unknown and where the variances are equal but otherwise unknown. Suppose we collect a sample of size n_X from X with sample mean $\bar{X}$ and sample variance s_Y^2, and suppose we collect a sample of size n_Y from Y with sample mean $\bar{Y}$ and sample variance s_Y^2.

(*i*) (10 points) Show that

$$Z := \frac{(\bar{X} - \bar{Y}) - (\mu_X - \mu_Y)}{\sigma\sqrt{1/n_X + 1/n_Y}} \sim \text{Norm}(0, 1).$$

(*ii*) (5 points, extra credit) Show that

$$V := \frac{(n_X - 1)s_X^2}{\sigma^2} + \frac{(n_Y - 1)s_Y^2}{\sigma^2} \sim \chi^2(n_X + n_Y - 2).$$

(*iii*) (15 points) You may use the (kind of deep) fact that the Z and V from parts (*i*) and (*ii*) are independent. Show that

$$\left((\bar{X} - \bar{Y}) - t_0 s_P\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}, \quad (\bar{X} - \bar{Y}) + t_0 s_P\sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}\right)$$

where

$$s_P = \sqrt{\frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2}}$$

and

$$t_0 = t_{\alpha/2}(n_X + n_Y - 2)$$

is a confidence interval for $\mu_X - \mu_Y$ with confidence coefficient $1 - \alpha$.

Problem 9

In this problem, we will test the hypothesis that the air quality in San Francisco is different from the air quality in Los Angeles. Let X and Y denote the concentration of suspended particles in the air in San Francisco and Los Angeles, respectively. As in Problem 8, assume $X \sim \mathrm{Norm}(\mu_X, \sigma^2)$ and $Y \sim \mathrm{Norm}(\mu_Y, \sigma^2)$ are normal, where the means are unknown and where the variances are equal but otherwise unknown. Suppose we collect data with the following statistics:

$$\begin{array}{ccc} n_X = 5 & \bar{X} = 79 & s_X = 5 \\ n_Y = 10 & \bar{Y} = 84 & s_Y = 5. \end{array}$$

(*i*) (20 points) Using the confidence interval described in Problem 8 part (*iii*), perform a hypothesis test on whether $\mu_X \neq \mu_Y$ at confidence level $\alpha = 0.05$. You should refer to the table of normal and t-distribution values on the last page of this exam, and you may use that $\sqrt{3/10} \approx 11/20$

(*ii*) (5 points) Give a bound (as tight as possible using the table at the end of the exam) for the p-value of your test. You may use that $1/11 \approx 0.09$.

Table of normal and t-distribution values

The following table gives some common values of the cdf of some t-distributions. Since $t(\infty) = \text{Norm}(0,1)$ this also contains common values of the cdf of the unit normal distribution. Recall that $t_\alpha(r)$ is defined as $t_\alpha(r) = F_{t(r)}(1-\alpha)$, where $t(r)$ denotes the t-distribution with r degrees of freedom.

r	$t_{0.10}(r)$	$t_{0.05}(r)$	$t_{0.025}(r)$	$t_{0.01}(r)$	$t_{0.005}(r)$	$t_{0.001}(r)$
1	3.078	6.314	12.706	31.821	63.657	318.31
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.215
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.785
8	1.397	1.860	2.306	2.896	3.355	4.501
9	1.383	1.833	2.262	2.821	3.250	4.297
10	1.372	1.812	2.228	2.764	3.169	4.144
11	1.363	1.796	2.201	2.718	3.106	4.025
12	1.356	1.782	2.179	2.681	3.055	3.930
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
35	1.306	1.690	2.030	2.438	2.724	3.340
40	1.303	1.684	2.021	2.423	2.704	3.307
45	1.301	1.679	2.014	2.412	2.690	3.281
50	1.299	1.676	2.009	2.403	2.678	3.261
55	1.297	1.673	2.004	2.396	2.668	3.245
60	1.296	1.671	2.000	2.390	2.660	3.232
∞	1.282	1.645	1.960	2.326	2.576	3.090