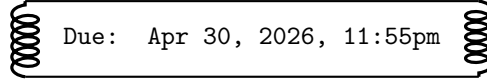


Homework 3



Instruction: Submit your answers electronically through Moodle. In Moodle, you must submit under the PDF Submission link with ONE PDF file containing answers to all remaining questions.

Question 1

A joint PDF of random variables X and Y is given by

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho \frac{(x-\mu_x)}{\sigma_x} \frac{(y-\mu_y)}{\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2} \right]},$$

where ρ denotes the correlation, with $-1 < \rho < 1$.

1. Show that if $\rho = 0$, then X and Y are independent, with each following a Gaussian PDF.
2. Show that for a general ρ but with $\mu_x = \mu_y = 0$ and $\sigma_x = \sigma_y = 1$:

(a) The expression can be simplified to

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)}.$$

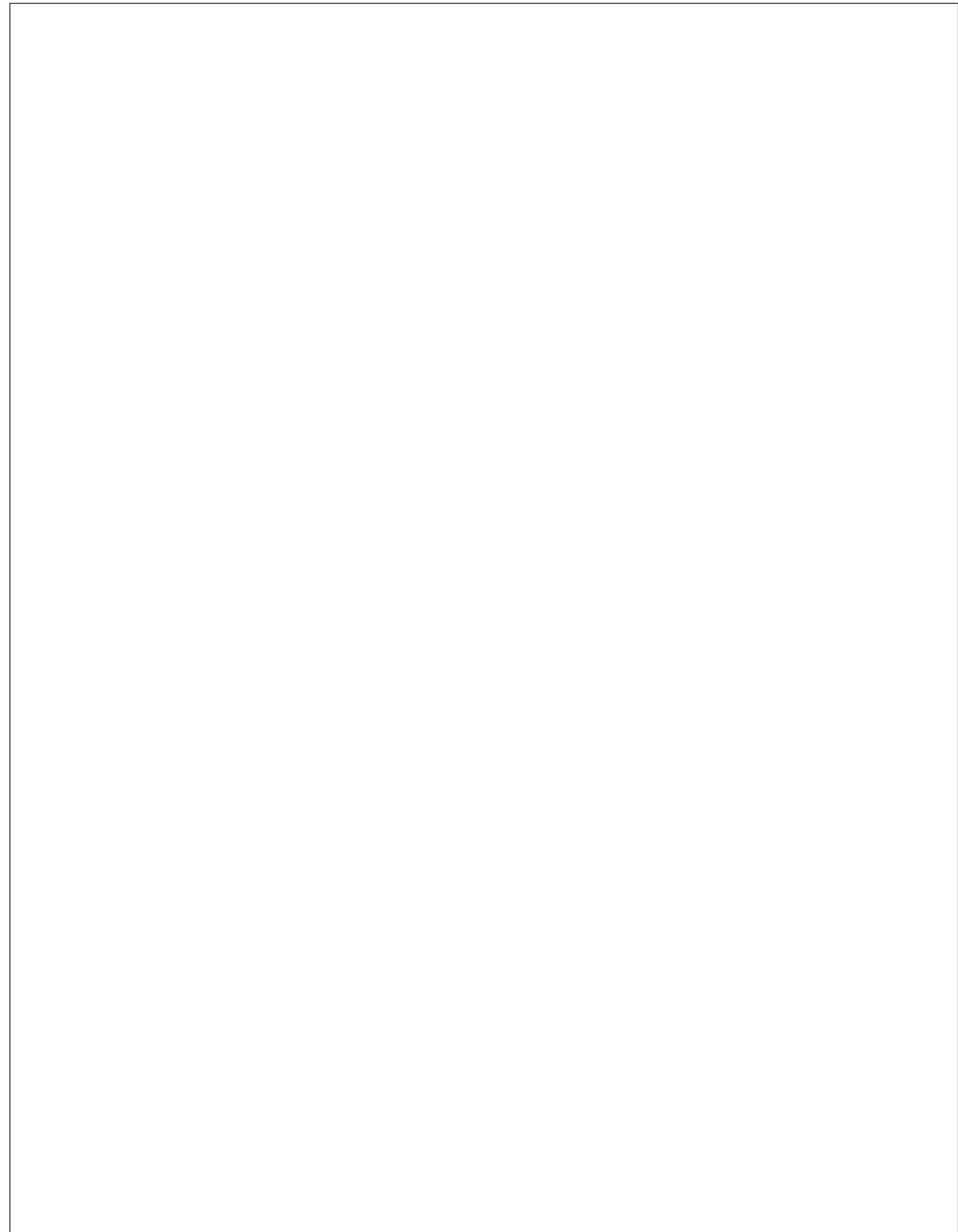
(b) Let $E = \frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}$. Show that we can also write

$$E = \frac{(y - \rho x)^2}{2(1-\rho^2)} + \frac{x^2}{2}.$$

(c) Let $u = y - \rho x$, and $\sigma^2 = 1 - \rho^2$. Show that the marginal PDF, given by

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy,$$

is also a Gaussian PDF.



Question 2

A certain continuous random variable X has probability density function (PDF) given by

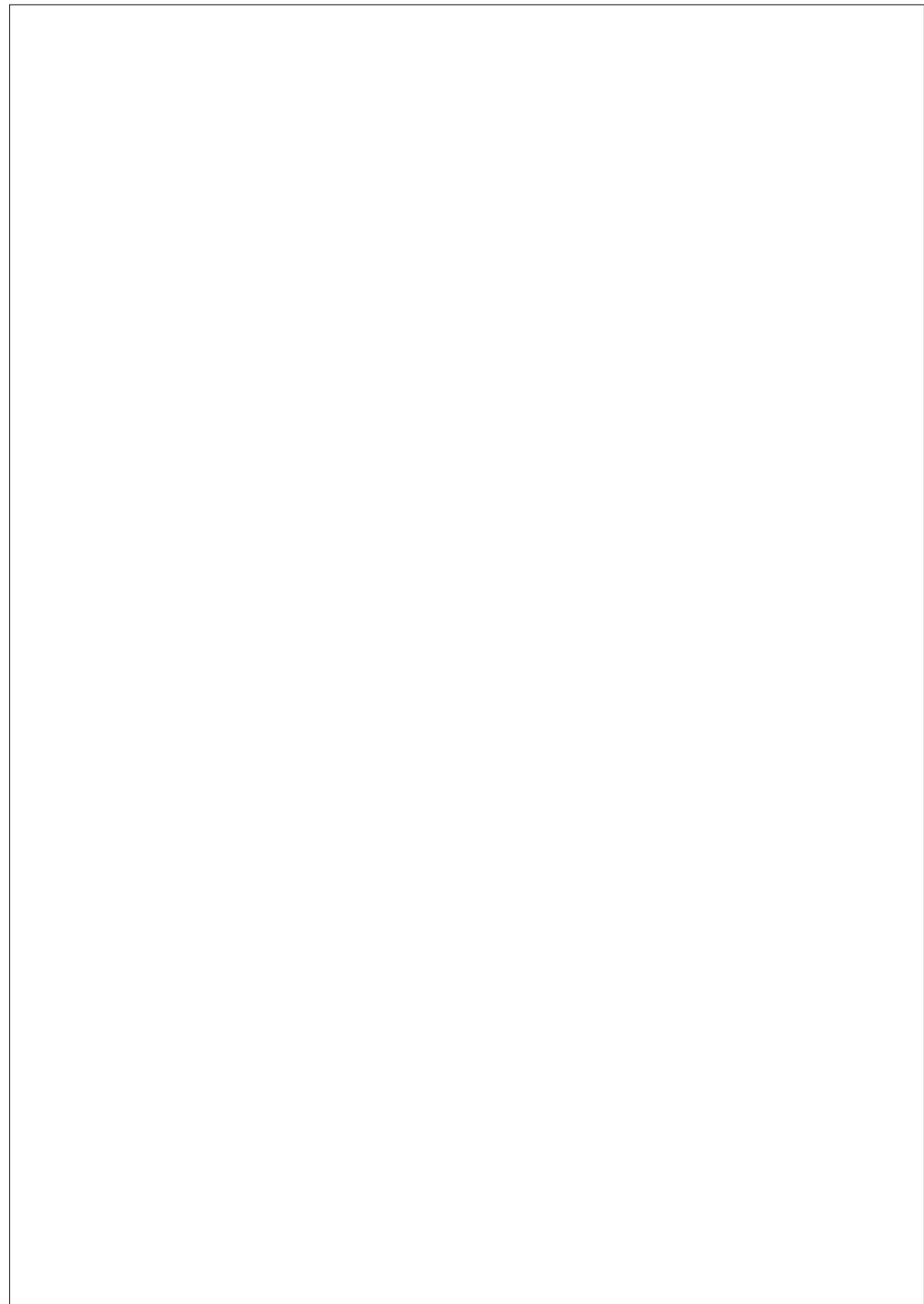
$$f_X(x) = \begin{cases} \frac{k}{\lambda^k} x^{k-1} e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

There are two parameters:

- A scale parameter, λ , which must be positive.
- A shape parameter, k , which must also be positive.

We made observations $X_1, \dots, X_n$, which are independent identically distributed and follow this PDF.

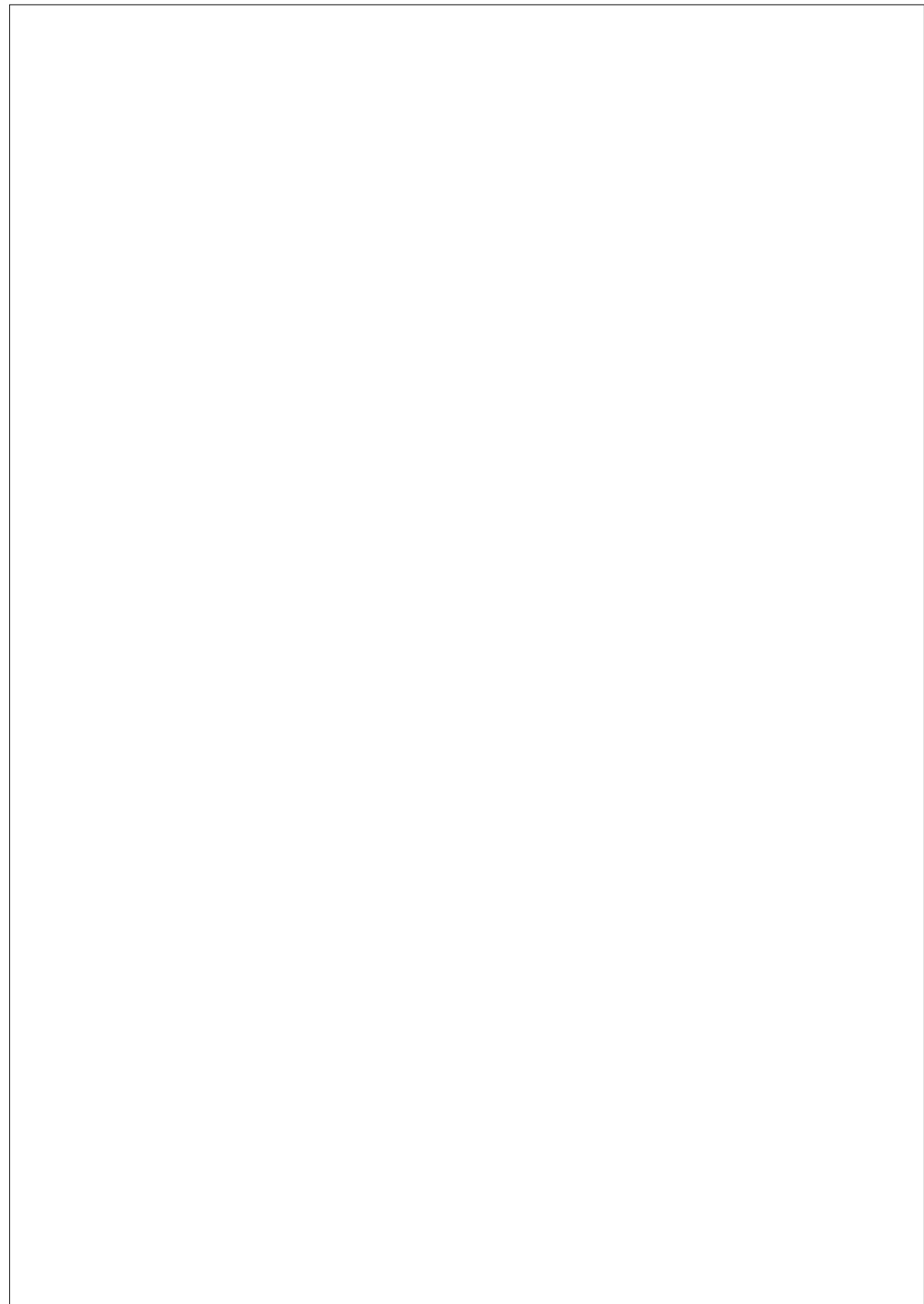
1. Suppose $k = 1$. Find the ML estimator of λ .
2. Suppose $k = 2$. Find the ML estimator of λ .
3. Find the ML estimator of λ for a general value of k .



Question 3

In an autonomous driving system, a radar module measures the distance to a moving vehicle. Suppose we collect n independent sensor readings $Y_1, Y_2, \dots, Y_n$. Each reading follows a normal distribution where the mean is proportional to an unknown velocity parameter θ and known time instances $x_1, x_2, \dots, x_n$. The sensor noise variance σ^2 is known. In other words, $Y_i \sim \mathcal{N}(\theta x_i, \sigma^2)$, where x_i and σ^2 are fixed and known, while θ must be estimated.

1. Find the maximum likelihood estimator $\hat{\theta}$.
2. Before seeing any sensor data, the car's predictive tracking algorithm assumes a normal distribution prior on θ , with specified mean equal to θ_0 and variance equal to τ^2 .
 - (a) Derive the posterior distribution of θ . (Hint: It is also a normal distribution, but you need to derive expressions for the mean θ_1 and variance σ_1^2 of this normal distribution.)
 - (b) Show that we can express the mean of the posterior distribution as a weighted sum of the prior mean θ_0 and the maximum likelihood estimator $\hat{\theta}$.
 - (c) Find the limit of the posterior distribution as $\tau^2 \rightarrow \infty$, and compare the estimates given by ML and MAP.



Question 4

A mobile application records the type of action performed by users when they open the app.

Each user action falls into one of three categories:

Browsing, Messaging, Shopping.

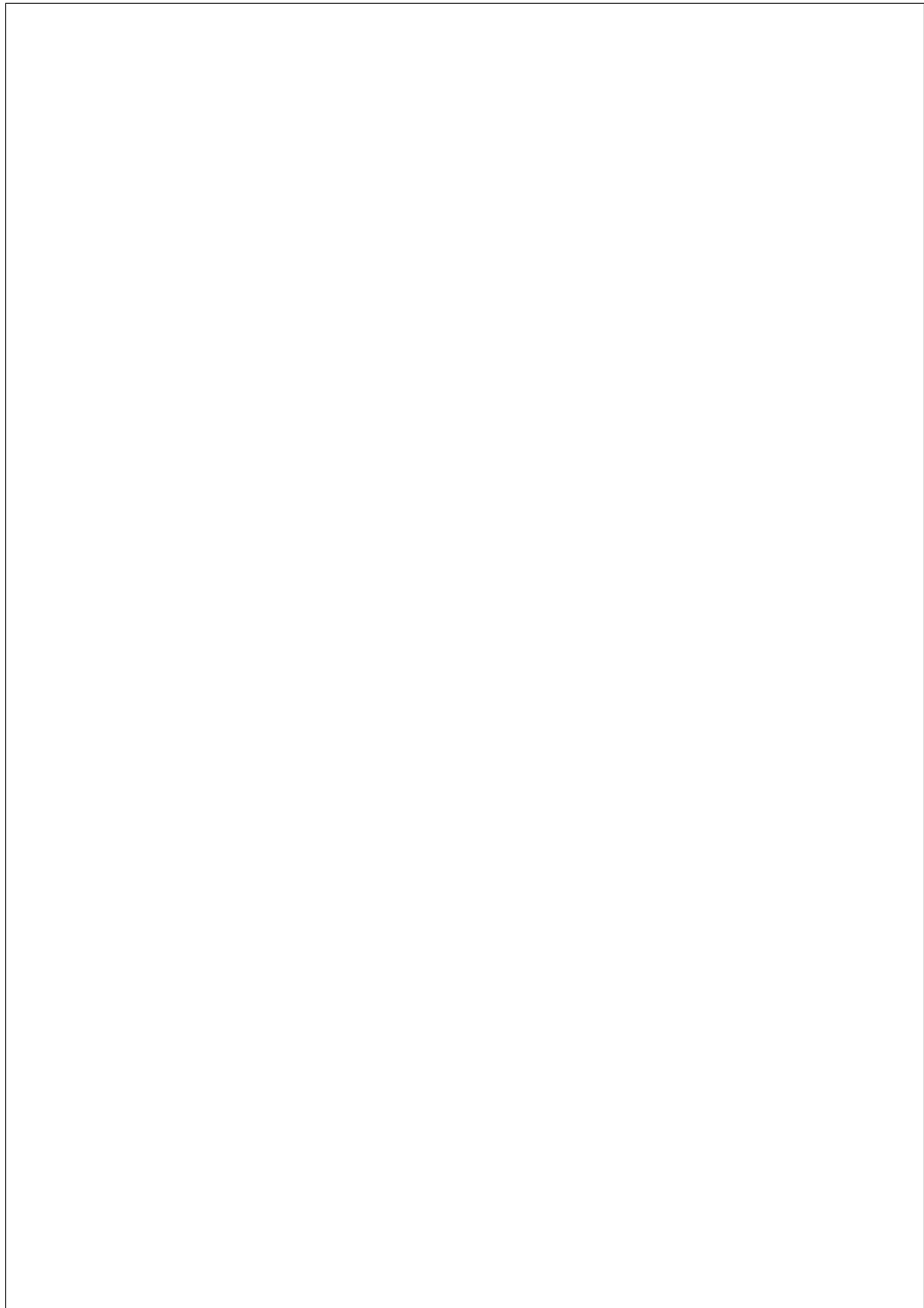
Let the probabilities of the three actions be θ_1 , θ_2 , and θ_3 , respectively, and assume that none of them is zero. Note also that $\theta_1 + \theta_2 + \theta_3 = 1$.

Before observing any data, an engineering intern, equipped with knowledge of probabilistic systems analysis, models the prior belief of the parameters by a Dirichlet distribution

$$(\theta_1, \theta_2, \theta_3) \sim \text{Dir}(\alpha_1, \alpha_2, \alpha_3).$$

The app then records n user sessions independently. Suppose that the observations show k_1 browsing actions, k_2 messaging actions, and k_3 shopping actions, with $k_1 + k_2 + k_3 = n$.

1. Write down the likelihood function of $(\theta_1, \theta_2, \theta_3)$.
2. Find the maximum likelihood estimators for $\theta_1, \theta_2, \theta_3$.
3. Derive the posterior distribution of $(\theta_1, \theta_2, \theta_3)$ and find the MAP estimators.



Question 5

A certain random variable X is modeled as an exponential distribution, with $X \sim \text{Exp}(\theta)$, i.e.,

$$f_X(x) = \begin{cases} \theta e^{-\theta x} & x \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

where θ is an unknown to be estimated from some observation data. We assume that θ follows a Gamma distribution denoted as $\Theta \sim \text{Gamma}(\alpha, \beta)$, given by

$$f_{\Theta}(\theta; \alpha, \beta) = \frac{\beta^{\alpha} \theta^{\alpha-1} e^{-\beta \theta}}{\Gamma(\alpha)},$$

where α and β are parameters of the distribution, and $\Gamma(\cdot)$ is the Gamma function. It is known that the mode of the Gamma distribution (with $\alpha \geq 1$) is given by $(\alpha - 1)/\beta$.

Assume we make a series of observations, X_i for $1 \leq i \leq n$, which are independent identically distributed, following the same exponential distribution.

1. Show that the posterior is given by

$$f_{\theta | X_1, \dots, X_n}(\theta | x_1, \dots, x_n) = c \cdot e^{-(\beta+s)\theta} \cdot \theta^{(\alpha+n)-1},$$

where $s = \sum_{i=1}^n x_i$, and c is a certain constant.

2. Express the posterior in terms of a Gamma distribution.
3. Find the MAP estimate of θ .
4. What happens to the MAP estimate when n is large?

