This is a take home midterm. You can use your notes, my online notes on canvas and the text book. You are supposed to work on your own text without external help. I'll be available to answer question in person or via email. Please, write clealy and legibly and take a readable scan before uploading.

Name (print): ______________________________

Question:	1	2	3	Total
Points:	30	20	50	100
Score:				

Question 1 .... *30 point*

Let X_i, $i = 1, \ldots, N$ be a random sample where the X_i are uniformly distributed in $[0, \theta]$, that is

$$f(x|\theta) = \frac{1}{\theta} \qquad \text{for} \qquad 0 \le x \le \theta\,.$$

(a) (10 points) Consider the family of distributions

$$g(\theta) = \frac{\alpha A^{\alpha}}{\theta^{\alpha+1}} \qquad \text{for} \qquad \theta \ge A$$

where α and A are hyperparameters. The distribution g is called a Pareto distribution with scale A and shape α. Show that this is a conjugate family of prior distribution for the sample $\mathbf{X}$. Assuming that the hyperparamter for the prior distribution are A_0 and α_0, find the hyperparamter α_1 and A_1 for the posterior distribution.

Solution: We get

$$g(\theta|\mathbf{x}) \propto \frac{1}{\theta^{\alpha_0+1}} \left(\frac{1}{\theta}\right)^N$$

for $\theta > \max(A_0, x_1, \ldots, x_N)$. Thus we have

$$g(\theta|\mathbf{x}) = \frac{(\alpha_0 + N) \max(A_0, x_1, \ldots, x_N)^{\alpha_0+N}}{\theta^{\alpha_0+N+1}}$$

so that

$$\alpha_1 = \alpha_0 + N \qquad\qquad A_1 = \max(A_0, x_1, \ldots, x_N)\,.$$

(b) (10 points) Find the Bayes estimator $\hat{\theta}_q(\mathbf{X})$ for the square error loss function and $\hat{\theta}_a(\mathbf{X})$ for the absolute error loss function.

Solution: We have

$$\hat{\theta}_q(\mathbf{x}) = \mathbb{E}(\theta|\mathbf{x}) = \int_{A_1}^{\infty} \theta \frac{\alpha_1 A_1^{\alpha_1}}{\theta^{\alpha_1+1}} = \frac{\alpha_1}{\alpha_1 - 1} A_1 = \frac{\alpha_0 + N}{\alpha_0 + N - 1} \max(A_0, x_1, \ldots, x_N)$$

On the other hand, $\hat{\theta}_a(\mathbf{x})$ is the median of $g(\theta|\mathbf{x})$, that is

$$\int_{A_1}^{\hat{\theta}_a} \frac{\alpha_1 A_1^{\alpha}}{\theta^{\alpha+1}} = \left(\frac{A_1}{\hat{\theta}_a}\right)^{\alpha_1} = \frac{1}{2}$$

and we get

$$\hat{\theta}_a = 2^{\frac{1}{\alpha_1}} A_1 = 2^{\frac{1}{\alpha_0+N}} \max(A_0, x_1, \ldots, x_N)$$

(c) (10 points) Show that $\hat{\theta}_q(\mathbf{X})$ and $\hat{\theta}_a(\mathbf{X})$ found above are consistent estimator, under suitable condition on A_0.(**Hint**: try first with $A_0 = 0$ and observe that if $Z_N \to_p \mu$ and $a_N \to 1$ as $N \to \infty$ then $a_N Z_N \to_p \mu$ as $N \to \infty$.)

Solution: Since both $\frac{\alpha_0+N}{\alpha_0+N-1}$ and $2^{\frac{1}{\alpha_0+N}}$ converge to 1 when $N \to \infty$, it is enough to discuss whether $Y_N = \max(A_0, X_1, \dots, X_N)$ converges in probability to θ_0, where θ_0 is the true value of θ. If $A_0 < \theta_0$,

$$\mathbb{P}(Y_N \le y) = \begin{cases} 0 & y < A_0 \\ \left(\frac{y}{\theta_0}\right)^N & A_0 \le y < \theta_0 \\ 1 & y \ge \theta_0 \end{cases}$$

thus, if $\delta > 0$,

$$\mathbb{P}(|Y_N - \theta_0| \ge \delta) = \mathbb{P}(Y_N \ge \theta_0 + \delta) + \mathbb{P}(Y_N \le \theta_0 - \delta) = \left(\frac{\max(A_0, \theta_0 - \delta)}{\theta_0}\right)^N .$$

Since $\theta_0 > A_0, \theta_0 - \delta$ we get $\mathbb{P}(|Y_N - \theta_0| \ge \delta) \to_{N\to\infty} 0$.
On the other hand if $A_0 \ge \theta_0$ then $Y = A_0$ with probability 1. Thus $\hat{\theta}_q(\mathbf{X})$ and $\hat{\theta}_a(\mathbf{X})$ are consisten if and only if $A_0 \le \theta_0$.

Question 2 .... *20 point*

Let X_i, $i = 1, \ldots, N$ be a random samble from a Bernoulli distribution with parameter 0.5, that is

$$p(0) = p(1) = 0.5\,.$$

Let

$$Y_N = \left(\prod_{i=1}^{N}(a + bX_i)\right)^{\frac{1}{N}}$$

where $a, b > 0$.

(a) (10 points) Use the Law of large number to show that Y_N converges in probability to a constant ν and find ν. (**Hint**: consider the r.v. $Z_N = \log(Y_N)$)

Solution: Observe that

$$Z_N = \frac{1}{N}\log(Y_N) = \frac{1}{N}\sum_{i=1}^{N}\log(a + bX_i) \to_p \mathbb{E}(\log(a + bX_i)) =$$
$$\frac{1}{2}\log(a) + \frac{1}{2}\log(a + b) = \log\left(\sqrt{a(a + b)}\right) = \mu$$

Since $Y_N = e^{Z_N}$ and the exponential is a continuous function we get

$$Y_N \to_p \sqrt{a(a + b)}$$

and $\nu = \sqrt{a(a + b)}$

(b) (10 points) Use the central limit theorem and the Delta method to find an approximate distribution for

Solution: Considering that

$$\mathrm{var}(\log(a + bX_i)) = \frac{1}{2}\log(a)^2 + \frac{1}{2}\log(a + b)^2 - \mu^2 = \log\left(\sqrt{\frac{a + b}{a}}\right)^2 = \sigma^2$$

the CLT tells us that

$$\frac{\sqrt{N}}{\sigma}(Z_N - \mu) \to_d \mathcal{N}(0, 1)\,.$$

Since $Y_N = e^{Z_N}$, from the Delta method we get

$$\frac{\sqrt{N}}{\sigma e^{\mu}}(Y_N - e^{\mu}) \to_d \mathcal{N}(0, 1)$$

so that Y_N has approximately the distribution of a Normal r.v. with average ν and variance $\nu^2\sigma^2$.

Question 3 .... *50 point*

Let X_i, $i = 1, \dots, N$ be a family of Bernoulli r.v. such that, for every $\mathbf{x} \in \{0,1\}^N$ and every i,

$$\mathbb{P}(X_{i+1} = x_{i+1} \,|\, X_i = x_i \,\&\, X_{i-1} = x_{i-1} \,\&\, \cdots \,\&\, X_1 = x_1) = \mathbb{P}(X_{i+1} = x_{i+1} \,|\, X_i = x_i)$$

and call

$$\mathbb{P}(X_{i+1} = x \,|\, X_i = y) = k(x, y)$$

where $x, y \in \{0,1\}$. Assume moreover that $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_1 = 1) = 1/2$. Thus the X_i form a simple Markov chain and k is the *transition kernel*. In the following we consider the case when

$$k(1 \,|\, 1) = k(0 \,|\, 0) = \theta \qquad\qquad k(1 \,|\, 0) = k(0 \,|\, 1) = (1 - \theta)\,.$$

(a) (10 points) Given an outcome $\mathbf{x} \in \{0,1\}^N$, find the probability $p(\mathbf{x}|\theta) = \mathbb{P}(\mathbf{X} = \mathbf{x} \,|\, \theta)$. (**Hint**: remember that $\mathbb{P}(A \cap B) = \mathbb{P}(A \,|\, B)\mathbb{P}(B)$)

Solution: Let $N_=(\mathbf{x})$ be the number of times x_i is equal to x_{i-1} and $N_{\neq}(\mathbf{x})$ be the number of times x_i is different from x_{i-1}. Clearly $N_=(\mathbf{x}) + N_{\neq}(\mathbf{x}) = N - 1$. The probablity of $\mathbf{x}$ is given by

$$p(\mathbf{x}|\theta) = \mathbb{P}(X_1 = x_1) \prod_{i=1}^{N-1} \mathbb{P}(X_{i+1} = x_{i+1} \,|\, X_i = x_i) = \frac{1}{2} \prod_{i=1}^{N-1} k(x_i, x_{i-1}) = \frac{1}{2}\theta^{N_=(\mathbf{x})}(1-\theta)^{N_{\neq}(\mathbf{x})}$$

(b) (10 points) Show that, if $\theta = 1/2$, the X_i represent the flipping of a fair coin. That is you have to show that $\mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = 1/2$ and that X_i is independent from X_j, for $i \neq j$.

Solution: If $\theta = \frac{1}{2}$ then $p(\mathbf{x}|\theta) = 2^{-N}$ for every $\mathbf{x}$, exactly what happens for the flipping of a fair coin.

(c) (10 points) Given the observed outcome $\mathbf{x}$ of $\mathbf{X}$, find the MLE estimate $\hat{\theta}_{MLE}(\mathbf{x})$ for θ.

Solution: To find $\hat{\theta}_{MLE}(\mathbf{x})$ we observe that the log likelihood function is

$$l(\theta) = N_{=}(\mathbf{x}) \log(\theta) + N_{\neq}(\mathbf{x}) \log(1-\theta)\,.$$

Differentiating we get

$$\frac{N_{=}(\mathbf{x})}{\theta} - \frac{N_{\neq}(\mathbf{x})}{1-\theta} = 0$$

that gives

$$\hat{\theta}_{MLE}(\mathbf{x}) = \frac{N_{=}(\mathbf{x})}{N-1}\,.$$

(d) (10 points) Show the the family of Beta distribution with hyperparameters α and β is a conjugate family of prior distribution. Find the updating rule for α and β. Find the squared error loss function estimator for θ.

Solution: If the prior is given by a Beta distribution with hyperparameters α_0 and β_0 then we have

$$g(\theta) \propto \theta^{\alpha_0-1}(1-\theta)^{\beta_0-1}$$

so that

$$g(\theta|\mathbf{x}) \propto \theta^{\alpha_0+N_{=}(\mathbf{x})-1}(1-\theta)^{\beta_0+N_{\neq}(\mathbf{x})-1}$$

so that the posterior is a Beta distribution with hyperparameters $\alpha_0 + N_{=}(\mathbf{x})$ and $\beta_0 + N_{\neq}(\mathbf{x})$. Wer know that the squared error loss function estimator for $\hat{\theta}_q(\mathbf{x})$ is the expected value of θ under the posterior distribution so that we get

$$\hat{\theta}_q(\mathbf{x}) = \frac{\alpha_0 + N_{=}(\mathbf{x})}{\alpha_0 + \beta_0 + N - 1}\,.$$

We know that the

(e) (10 points) Can you find a Method of Moment moment estimate for θ?

Solution: Observe hat

$$\mathbb{P}(X_1 = 0) = \mathbb{P}(X_0 = 0)k(0\,|\,0) + \mathbb{P}(X_0 = 1)k(0\,|\,1) = \frac{1}{2}$$

so that

$$\mathbb{P}(X_1 = 1) = \frac{1}{2}\,.$$

Iterating this argument we get that

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = 0) = \frac{1}{2}$$

for every i. Thus $\mathbb{E}(X_i) = \frac{1}{2}$ for every i and it is not possible to use it to estimate θ.