Question-1 [6 points]: . In a study, researchers wanted to compare two types of sodas in terms of the volume of foams that they produce when poured in. They took 12 cans (all of the same size) of each type of sodas (total 24 cans). The table below summarizes their findings.

	Observations	Sample mean	Sample SD
Soda 1 (X)	16.69,16.99, … ,15.73	16.15	1.65
Soda 2 (Y)	19.75,17.05, … ,18.76	19.12	2.49

a. [2 points] Test $H_0: \sigma_x^2 = \sigma_y^2$ vs. $H_a: \sigma_x^2 \neq \sigma_y^2$ at $\alpha = 0.1$. What conclusion do you make?

雅婧 Gong 9987

b.[4 points] They want to see if the data provide evidence that the two sodas are different in terms of average foam volume. By assuming the population variances to be unequal, test $H_0: \mu_x = \mu_y$ vs. $H_a: \mu_x \neq \mu_y$ at $\alpha = 0.1$. What conclusion do you make?

雅婧 雅婧 雅婧 雅婧 雅婧 雅婧 雅婧 雅婧 雅婧 雅婧 雅婧 雅婧 雅婧 雅婧 雅婧

Gong 9987

Question-2 [7 points]: Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. with the following density function

$$f(x) = \frac{1}{\beta} e^{-\frac{(x-5)}{\beta}}; x \geq 5; \beta > 0$$

You are also told $E[X_i] = \beta + 5$.

a.[2 points] Using factorization theorem, find a sufficient statistic for β.

b. [3 points] Show that $E\left[S(\beta \mid X_1, X_2, \ldots X_n)|_{\beta=\beta_0}\right] = 0$, where β_0 is the true value of β. (this question is NOT asking for the generalized proof of "expectation of score is zero" that we did in the lecture).

22122964079

c. [2 points] Calculate the Cramer-Rao lower bound (CRLB).

雅婧 — Gong 9987

22122964079

Question-3 [7 points]: Suppose a market research company is interested in the proportion of households (π) in Toronto that uses Rogers as the internet service provider(ISP).

They want to test $H_0: \pi = 0.3$ vs $H_1: \pi \neq 0.3$
They randomly contact 150 households and find 60 of them using Rogers as the ISP.
a.[2 points] Construct a 95% confidence interval for π.

b. [1 point] At 5% level of significance ($\alpha = 0.05$), test the given hypothesis and comment on the conclusion that you make.

雅婧 Gong 9987

c. [4 points] Using $\frac{\bar{X}-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$ as the test statistic and by defining the rejection region under the null hypothesis, calculate the power of the test, $H_0: \pi = 0.3$ vs $H_1: \pi = 0.4$ for sample size, $n = 150$ and $\alpha = 0.05$

22122964079

Question-4 [6 points]: Suppose $Y_1, Y_2, \dots, Y_n$ are independently drawn from the following pdf

$$f(y) = \frac{1}{\theta} e^{-\frac{1}{\theta} y}; y > 0;\ \theta > 0$$

where θ is the unknown parameter.

a. [4 points] Construct a Wald test for testing $H_0: \theta = \theta_0$ vs $H_1: \theta \neq \theta_0$ at a given α. (constructing a test means providing the form of the test statistic, appropriate distribution and the rejection region)

雅婧 — Gong 9987

b.[2 points] Suppose these following observations are independently drawn from the above pdf

$$1.35, 1.95, 0.72, 0.79, 1.96, 1.81, 0.17, 0.72, 1.54, 0.18$$

with $\sum y_i = 11.19$.
Test $H_0: \theta = 2$ vs $H_1: \theta \neq 2$ at $\alpha = 0.05$ using the Wald test that you constructed in part(a). If you could not answer part(a), use a likelihood ratio test.

22122964079

Question-5 [6 points]: Suppose $x_1, x_2, \ldots, x_n$ are independently drawn from the following pmf.

$$P[X = x \mid p] = \binom{x-1}{4} p^5 (1-p)^{x-5}\ ;\ \text{where } x = 5,6,7,\ldots$$

Suppose, p follows a Beta (α, β) distribution with pdf,

$$f(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}; 0 < p < 1$$

a. [3 points] Find the posterior distribution of p. (name the distribution and provide the parameters).

雅婧 Gong 9987

b.[1.5 points] Comment on the type of this prior and propose a non-informative prior.

b. [1.5 points] What is the posterior mean? Under a non-informative prior calculate the posterior Mode

Gong 9987

Question-6 [6 points]: Suppose

$$(X_1, X_2, \ldots, X_n) \overset{iid}{\sim} \text{Bernoulli}\,(\pi_x) \text{ and } (Y_1, Y_2, \ldots, Y_m) \overset{\text{iid}}{\sim} \text{Bernoulli}\,(\pi_y)$$

You are also told that X and Y are independent.

a.[4 points] Construct a likelihood ratio test (LRT) for testing the hypothesis, $H_0: \pi_x = \pi_y$ with an alternative hypothesis, $H_a: \pi_x \neq \pi_y$ at significance level α. (This question is asking for three things: the test statistic, corresponding distribution and the rejection region). Simplify your test statistic as much as possible.

雅婧 雅婧 雅婧
雅婧 雅婧 雅婧
雅婧 雅婧 雅婧
雅婧 雅婧 雅婧
雅婧 雅婧 雅婧
—Gong 9987

b.[2 points] Suppose these following two set of samples are believed to be from a Bernoulli (π_x) and a Bernoulli (π_y) distributions.

X	1,1,1,1,0,1,0,1,1,0	$\sum_{i=1}^{10} x_i = 7$
Y	0,1,1,0,0,1,0,0,0,1	$\sum_{i=1}^{10} y_i = 4$

Using the LRT derived in part (a), test $H_0: \pi_x = \pi_y$ vs. $H_a: \pi_x \neq \pi_y$ at $\alpha = 0.05$

雅婧 — Gong 9987

Question-7 [7 points]: Suppose you are given a dataset with 5 observations. The first two columns of the following table represent the given data:

x_i	y_i	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
-1	5	4	9	6
0	7	1	1	1
1	8	0	0	0
2	9	1	1	1
3	11	4	9	6
Sum: 5	40	10	20	14

a. [1 point] Write down the complete likelihood function for the given data under the regression model $Y \mid X = x \sim N(\beta_1 + \beta_2 x, \sigma^2)$.

UTSC 导师：雅婧

22122964079

b. [1.5 points] Calculate an estimate of β_1 and β_2 and interpret their values.

c. [1.5 points] Calculate the value of the coefficient of determination and interpret the value.

22122964079

d. [1.5 points] Test (at $\alpha = 0.05$) the hypothesis, $H_0: \beta_2 = 2$ vs $H_a: \beta_2 \neq 2$. Write your complete conclusion.

e. [1.5 points] Suppose we create a new variable called W where $W = X/10$. We fit the regression model of $Y \mid W = w \sim N(\beta_1^* + \beta_2^* w, \sigma^2)$. Write down β_1^* and β_2^* in terms of β_1 and β_2.