1. Each part below contains an incomplete paragraph and some phrases that could complete it. Report the best choice from the given options and give a brief justification of your choice. \[6 \times 4 = 24 \text{ points}\]

(a) A count data \(X\) is modeled as \(X \sim f(x|\theta), \theta \in \{0, 1\}\) as in the table below

| \(\theta\) | \(f(1|\theta)\) | \(f(2|\theta)\) | \(f(3|\theta)\) |
| --- | --- | --- | --- |
| 0 | 0.9 | 0.05 | 0.05 |
| 1 | 0.09 | 0.055 | 0.855 |

In testing \(H_0 : \theta = 0\) vs. \(H_1 : \theta = 1\) the rule “reject \(H_0\) if & only if \(X = 2\)” gives

A. a 5% ML test  
B. a 5% non-ML test  
C. an ML test with size \(\neq 5\%\)  
D. a test that is neither 5% nor ML.

(b) Weekly food expenditures \(X_1, \ldots, X_n\) of \(n = 22\) Duke students are modeled as \(X_i \overset{iid}{\sim} \text{Normal}(\mu, \sigma^2)\). From observed data a 99% ML confidence interval for \(\mu\) is constructed to be \([120.9, 166.3]\). For testing \(H_0 : \mu \leq 150\) vs \(H_1 : \mu > 150\), the size 5% ML test

A. will reject \(H_0\)  
B. will accept \(H_0\)  
C. may reject or accept \(H_0\), we need more information to determine which.

(c) A sample of \(n = 30\) cheese slices were measured for their lactic acid concentrations and were found to have mean = 1.44 and standard deviation 0.30. To test whether the measurements are IID draws from a \(\text{Normal}(\mu, \sigma^2)\) for some \(\mu \in (-\infty, \infty), \sigma > 0\), we used the standard Pearson’s test with \(k = n/5 = 6\) equal probability bins under \(\text{Normal}(1.44, 0.30^2)\) given by

\((-\infty, 1.15], (1.15, 1.31], (1.31, 1.44], (1.44, 1.57], (1.57, 1.74], (1.74, \infty)\)

and with observed bin counts 5, 8, 2, 5, 4, 6. This leads to Pearson’s \(Q = 4\) and a p-value range = \([0.26, 0.55]\). Later it was discovered that one of the measurements was wrongly recorded as “2.01” when the correct record should have been “2.10”. This finding will affect calculations of

A. both \(Q\) and the p-value range  
B. neither \(Q\) nor the p-value range.

(d) In a three-way classification, 200 university students are simultaneously classified according to their college year (4 categories), major type (5 categories) and political affiliation (3 categories). To test \(H_0 : \text{“college year is independent of major type and political affiliation”}\), one could use a Pearson’s chi-squares test statistic \(Q\), whose distribution under \(H_0\) is approximately \(\chi^2(r)\) with

A. \(r = 21\)  
B. \(r = 24\)  
C. \(r = 42\)  
D. \(r = 59\)

(e) A pair of counts \(X = (X_1, X_2)\) are modeled as \(X_i \sim g(x_i|\theta), \theta \in \{0, 1\}\) where the pmfs \(g(x_i|\theta)\), defined over \(x_i \in \{0, 1\}\) (and \(g(x_i|\theta) = 0\) otherwise) are given as below:
| $\theta$ | $g(0|\theta)$ | $g(1|\theta)$ |
|---------|--------------|--------------|
| 0       | $\frac{1}{5}$ | $\frac{4}{5}$ |
| 1       | $\frac{1}{2}$ | $\frac{1}{2}$ |

What is the size of the ML test: “reject $H_0$ if & only if $LR(X) > 1$”? 
A. $\frac{1}{25}$  B. $\frac{1}{5}$  C. $\frac{9}{25}$  D. $\frac{2}{5}$

(f) Consider the same data and model as in part (e) above and suppose $\theta$ is assigned the discrete uniform pmf $\xi(0) = \xi(1) = 1/2$. What is the posterior probability that $\theta = 0$ if we observe $X_1 = 0$ and $X_2 = 1$?
A. 0.18  B. 0.29  C. 0.39  D. 0.50

2. The NSW study had two arms, a treatment arm of $n = 185$ subjects who received job training and a control arm of $m = 260$ subjects who did not receive any training but were otherwise similar to subjects in the treatment arm. We have data on both groups’ annual incomes in the base year 1975 and the follow-up year 1978.

Let $X_1, \ldots, X_n$ denote the income difference between follow-up and base years for treatment subjects, and $Y_1, \ldots, Y_m$ denote the same for the control group, modeled as: $X_i \overset{\text{iid}}{\sim} \text{Normal}(\mu_1, \sigma^2)$, $Y_j \overset{\text{iid}}{\sim} \text{Normal}(\mu_2, \sigma^2)$, $-\infty < \mu_1, \mu_2 < \infty$, $\sigma^2 > 0$. Observed data show: $\bar{X} = 4817.1$, $s_X = 8275.4$, $\bar{Y} = 3287.8$, $s_Y = 6059.8$ and therefore any $100(1-\alpha)\%$ ML confidence interval for $\mu_1 - \mu_2$ is given by $1529.2 \pm 679.5 \cdot z_{.443}(\alpha)$.

(a) What is the ML p-value for testing $H_0 : \mu_1 \leq \mu_2$ against $H_1 : \mu_1 > \mu_2$? [3 points]

(b) Report the 95% posterior credible interval for $\mu_1 - \mu_2$ under the reference prior $\xi(\mu_1, \mu_2, \sigma^2) = 1/\sigma^2$. [3 points]

(c) In the treatment group $U = 120$ subjects reported a substantial increase in income (defined as 1978 income being at least 25% more than 1975 income). $V = 146$ reported the same in the control group. Consider the model $U \sim \text{Binomial}(n, p_1)$, $V \sim \text{Binomial}(m, p_2)$, $U$ and $V$ conditionally independent given $p_1, p_2 \in [0,1]$. Under what joint prior $\xi(p_1, p_2)$ is the posterior equivalent to: $p_1 \sim \text{Beta}(121, 66)$, $p_2 \sim \text{Beta}(147, 115)$, $p_1, p_2$ independent? Justify. [3 points]

(d) In the setting of part (c), report a 95% posterior range for $p_1 - p_2$ and give a brief justification. [Hint: if $a, b$ are large then $\text{Beta}(a, b) \approx \text{Normal}(m, s^2)$ with $m = \frac{a}{a+b}$ and $s^2 = \frac{m(1-m)}{a+b+1}$] [4 points]

3. A sample of $n = 272$ eruption durations (in minutes) $X_1, \ldots, X_n$ of a geyser are modeled as $X_i \overset{\text{iid}}{\sim} \text{Uniform}(0, \theta)$, $\theta > 0$. By central limit theorem, whenever $X_i \overset{\text{iid}}{\sim} \text{Uniform}(0, \theta)$, the distribution of $\bar{X}/\theta$ is well approximated by $\text{Normal}(\frac{1}{2}, \frac{1}{12n})$. We want to test $H_0 : \theta \geq 6$ against $H_1 : \theta < 6$.

(a) Justify why it is reasonable in this case to reject $H_0$ for large values of the test statistic $T(X) = 6/\bar{X}$? [2 points]
(b) For any arbitrary $\alpha \in (0, 1)$ what is the size-$\alpha$ test based on $T(X)$. It is enough to precisely state the rejection rule, you do NOT have to rigorously prove that your test has size $\alpha$.

(c) The size-5% test based on $T(X)$ is: “reject $H_0$ if & only if $6/\bar{X} > 2.12$”. On the other hand, the 5% ML test is: “reject $H_0$ if & only if $6/X_{\text{max}} > 1.011$”. For observed data, $\bar{X} = 3.49$ and $X_{\text{max}} = 5.1$. What should you report about the validity of $H_0$? Give a brief, coherent justification, no proofs needed. [3 points]

4. Annual hurricane counts $X_1, \ldots, X_n$ from last $n = 10$ years (2004 through 2013) are modeled as $X_i \overset{\text{iid}}{\sim} \text{Poisson}(\mu)$, $\mu \in (0, \infty)$. Observed data shows $\bar{X} = 15.4$.

(a) Report the 95% ML confidence interval for $\mu$ (show work). [4 points]

(b) Report the 95% posterior credible interval for $\mu$ under the Jeffreys prior $\xi(\mu) = 1/\sqrt{\mu}$. [Hint: Gamma$(a, b) \approx \text{Normal}(a/b, a/b^2)$ if $a$ is large.] [4 points]

(c) Historical records drawn from 1984 through 2003 shows roughly 10 hurricanes per year and $\xi(\mu) = \text{Gamma}(200, 20)$ presents a good contextual prior for $\mu$. By conjugacy, the corresponding posterior pdf is $\xi(\mu|x) = \text{Gamma}(200+10\times15.4, 20+10) = \text{Gamma}(354, 30)$.

Consider a formal Bayes testing of $H_0 : \mu \leq 12$ (no substantial increase in hurricane activity) against $H_1 : \mu > 12$ (20%+ increase) under the above contextual prior. Which decision will you take if the loss calculations were given as follows?

<table>
<thead>
<tr>
<th>Truth</th>
<th>No increase</th>
<th>20%+ increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu \leq 12$</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>$\mu &gt; 12$</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

You may use the facts that the event “$\mu \leq 12$” receives 99.6% chance under $\text{Gamma}(200, 20)$ and 63.1% chance under $\text{Gamma}(354, 30)$. [4 points]

(d) It is known that 12 is below the left boundary of the 98% ML confidence interval for $\mu$ but is included in the 99.5% ML confidence interval. Should the size 1% ML test reject $H_0 : \mu \leq 12$ against $H_1 : \mu > 12$? [3 points]

5. 100 randomly chosen customers were emailed a web link to a survey questionnaire regarding a certain product they bought from an online store. Among these

- $X_1 = 25$ did not click on the link
- $X_2 = 21$ clicked on the link but did not submit the survey
- $X_3 = 24$ clicked & submitted the survey conveying dissatisfaction with product
- $X_4 = 30$ clicked & submitted the survey, conveying satisfaction with product.

Assume truthful reporting. We want to test $H_0$: “a third of the customers are dissatisfied with the product” under the following assumptions on response rates:

$$\frac{P(\text{click} \mid \text{dissatisfied})}{P(\text{click} \mid \text{satisfied})} = 2, \quad \frac{P(\text{submit} \mid \text{click, dissatisfied})}{P(\text{submit} \mid \text{click, satisfied})} = 2$$
We model \( X = (X_1, X_2, X_3, X_4) \) by \( X \sim \text{Multinomial}(100, p) \), \( p = (p_1, p_2, p_3, p_4) \in \Delta_4 \).

(a) Show that under \( H_0 \) and the assumptions on response rates:

\[
\frac{p_3}{p_4} = \frac{P(\text{dissatisfied, click, submit})}{P(\text{satisfied, click, submit})} = 2
\]

(b) Let \( a = P(\text{click}) \) and \( b = P(\text{submit | click}) \). Argue that under the assumptions on the response rates, \( H_0 \) corresponds to the null subset

\[
\Delta^0_4 = \{(1 - a, a(1 - b), \frac{2}{3}ab, \frac{1}{3}ab) : a \in (0, 1), b \in (0, 1)\}.
\]

(c) Show that under the assumptions on the response rates, the restricted mle of \( p \) under \( H_0 \) is given by \( \hat{p}_0 = (0.25, 0.21, 0.36, 0.18) \). [Hint: For any two non-negative numbers \( r \) and \( s \) the function \( g(a) = a^r(1 - a)^s \) over \( a \in (0, 1) \) is maximized at \( \hat{a} = \frac{r}{r + s} \) ]

(d) Will you reject \( H_0 \) at 1% significance level?