Final Examination

Mth 136 = Sta 114

Monday, 2011 May 2, 7:00 – 10:00pm

- This is a **closed book** exam— please put your books on the floor.
- You may use a calculator and **two pages** of your own notes.
  Do not share calculators or notes.
- **Show** your **work.** Neatness counts. $\boxed{\text{Boxing answers}}$ helps.
- **Simplify all expressions** for full credit. Numerical answers: **four significant digits** or fractions in lowest terms.
- Distribution & pdf/pmf tables and blank worksheet are attached.

Cheating on exams is a breach of trust with classmates and faculty that will not be tolerated. After completing the exam please acknowledge that you have adhered to the Duke Community Standard:

- I will not lie, cheat, or steal in my academic endeavors;
- I will conduct myself honorably in all my endeavors; and
- I will act if the Standard is compromised.

Please affirm the Community Standard with your signature below.

Signature: Li Ma

Print Name: Li Ma

1.	/20	6.	/20
2.	/20	7.	/20
3.	/20	8.	/20
4.	/20	9.	/20
5.	/20	10.	/20
Total:	/200		

Problem 1: Smokey Joe likes to cook outside on the grill. The thing he likes best is that, every now and then the coals burst into violent flame and totally incinerate whatever he's cooking. Each day of the week this happens independently, always with the same probability p (which might be different one week than another). It's very exciting.

Each day, the number of hamburgers Smokey cooks before the bonfire has a geometric probability distribution, with pmf

$$\mathsf{P}[X_i = x \mid p] = p(1-p)^x, \quad x = 0, 1, 2, \cdots$$

a) Find a sufficient statistic $T_n(\mathbf{x})$ for n observations $\mathbf{x} = \{x_1, ..., x_n\}$ from this distribution, and explain why T_n is sufficient.

$T_n(\mathbf{x}) = \underline{\quad \sum_{i=1}^{n} X_i \quad}$

$$L(p) = f_n(\underline{x}) = \prod_{i=1}^{n} p(1-p)^{x_i} = p^n \cdot (1-p)^{\sum_{i=1}^{n} x_i}$$

b) Find the Maximum Likelihood Estimate (MLE) $\hat{p}(\mathbf{x})$ for p, if we observe the data set $\mathbf{x} = \{7, 5, 9, 6, 3\}$. Show your work— *derive* the result, don't just write down the answer. Answer is numeric.

$\hat{p} = \underline{\quad \frac{n}{n + \sum_{i=1}^{n} X_i} \quad} = \frac{5}{5 + 7 + 5 + 9 + 6 + 3} \approx .167.$

$$L(p) = p^n \cdot (1-p)^{\sum_{i=1}^{n} x_i}$$

$$\log L(p) = n \log p + \left(\sum_{i=1}^{n} x_i\right) \cdot \log(1-p)$$

$$\frac{d}{dp} \log L(p) = \frac{n}{p} + \left(\sum_{i=1}^{n} x_i\right) \cdot \frac{(-1)}{1-p} = 0$$

$$\Rightarrow n(1-p) = p \cdot \sum_{i=1}^{n} X_i \Rightarrow \hat{p} = \frac{n}{n + \sum_{i=1}^{n} X_i}$$

$$\frac{d^2}{dp^2} \log L(p) = -\frac{n}{p^2} - \left(\sum_{i=1}^{n} X_i\right) \cdot \frac{1}{(1-p)^2} < 0. \text{ So it's MLE.}$$

Problem 1 (cont):

c) For the same data set $\mathbf{x} = \{7, 5, 9, 6, 3\}$, with a uniform prior distribution $\pi(p) = \mathbf{1}_{\{0<p<1\}}$, find the posterior mean and variance:[1]

$\mathsf{E}[p \mid \mathbf{x}] =$ $\frac{6}{37}$ $\qquad$ $\mathsf{Var}[p \mid \mathbf{x}] =$ $.0036$.

$$\pi(p|\underline{x}) \propto \pi(p)\cdot f_n(\underline{x}) = p^n(1-p)^{\sum_{i=1}^n x_i} \quad \text{for } 0<p<1$$

$$= 0 \quad \text{o.w.}$$

$$\text{So } \pi(p|\underline{x}) \sim Beta(n+1, 1+\sum_{i=1}^n x_i)$$

$$\sim Beta(6, 31).$$

$$E(p|\underline{x}) = \frac{6}{6+31} = \frac{6}{37} \qquad Var(p|\underline{x}) = \frac{6\times 31}{(6+31)^2(6+31+1)} \approx .0036$$

d) With the same uniform prior distribution $\pi(p) = \mathbf{1}_{\{0<p<1\}}$, find the posterior probability that $p \le \frac{1}{2}$ for a different week— a really bad one with seven mishaps in a row, $\mathbf{x} = \{0, 0, 0, 0, 0, 0, 0\}$.

$\mathsf{P}[p \le \frac{1}{2} \mid \mathbf{x}] =$ $\frac{1}{256}$.

$$\pi(p|\underline{x}) \propto p^7\cdot(1-p)^0 = p^7$$

$$\pi(p|\underline{x}) = 8p^7$$

$$\text{So } P(p\le \tfrac{1}{2}|\underline{x}) = \int_0^{\frac{1}{2}} 8p^7\,dp = p^8\Big|_0^{\frac{1}{2}} = \left(\tfrac{1}{2}\right)^8 = \frac{1}{256}.$$

[1]Hint: No integration is needed for part c). Answer is numeric.

Problem 2: Let X have the Poisson distribution $\mathsf{Po}(\lambda)$ with pmf

$$f(x \mid \lambda) = \frac{\lambda^x}{x!} e^{-\lambda}, \qquad x = 0, 1, 2, \cdots .$$

~~a)~~ Write the pmf for X in natural exponential family form

$$f(x \mid \eta) = e^{\eta \cdot T(x) - A(\eta)}\, h(x)$$

$f(x|\lambda) = e^{x \log \lambda - \lambda} \cdot \frac{1}{x!}$

by identifying the:

Natural Parameter $\eta(\lambda) =$ $\log \lambda$

Sufficient Statistic $T(x) =$ x

Normalizing Constant $A(\eta) =$ $\lambda = e^{\eta}$

Data Function $h(x) =$ $\frac{1}{x!}$

b) Find the Fisher Information[2] $I_\eta(\eta)$ for the natural parameter[3]:

$I_\eta(\eta) =$ $A''(\eta) = e^{\eta}$.

$$\log L(\eta) = \log f(x|\eta) = \eta \cdot T(x) - A(\eta) + \log h(x).$$

$$\frac{d \log L(\eta)}{d\eta} = T(x) - A'(\eta).$$

$$\frac{d^2}{d\eta^2} \log L(\eta) = -A''(\eta).$$

So $I_\eta(\eta) = E\left(-\frac{d^2}{d\eta^2} \log L(\eta)\right) = A''(\eta) = e^{\eta} (= \lambda).$

[2] In parts b) and c) of this problem you're asked to find two different Fisher information functions— one for X's distribution when parametrized by λ, and one for the same distribution parametrized by η. The subscript (λ or η) is just a label so we don't mix them up. Only one observation is used in parts a, b, c.

[3] Hint: Find $A(\eta)$ first! You'll need it!

Problem 2 (cont):

c) Find the Fisher Information $I_\lambda(\lambda)$ for the usual (mean) parameter:
$I_\lambda(\lambda) =$ $\frac{1}{\lambda}$.

Two ways to get this:

① $\log L(\lambda) = x\log\lambda - \lambda - \log(x!)$

$\frac{d}{d\lambda}\log L(\lambda) = \frac{x}{\lambda} - 1$

$\frac{d^2}{d\lambda^2}\log L(\lambda) = -\frac{x}{\lambda^2}$ $\quad E\left(-\frac{d^2}{d\lambda^2}\log L(\lambda)\right) = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$

② $I_\lambda(\lambda) = E\left(\left(\frac{d\log L(\lambda)}{d\lambda}\right)^2\right)$

$= E\left(\left(\frac{d\log L(\eta)}{d\eta}\cdot\frac{d\eta}{d\lambda}\right)^2\right)$

$= I_\eta(\eta)\cdot\left|\frac{d\eta}{d\lambda}\right|^2$

$= e^\eta\cdot\left(\frac{1}{\lambda}\right)^2$

$= \lambda\cdot\left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda}$.

d) The MLE $\hat\lambda(\mathbf{x}) = \bar{x}_n$ of the Poisson mean parameter λ for a sample $\mathbf{x} = \{X_1, ..., X_n\}$ of size n is unbiased. What does the Cramér-Raô (Information) Inequality say about its squared-error risk, $\mathsf{E}[(\hat\lambda(\mathbf{x}) - \lambda)^2 \mid \lambda]$?

$$E\left((\hat\lambda(x)-\lambda)^2 \mid \lambda\right) > \frac{1}{n I_\lambda(\lambda)} = \frac{\lambda}{n}.$$

e) Find the MLE $\hat\eta_n(\mathbf{x})$ of the natural parameter[4] η.
For just 1 pt: Do you expect that $\hat\eta_n$ is unbiased too? Why?
$\hat\eta_n(\mathbf{x}) =$ $\log\hat\lambda(\underline{x}) = \log\bar{x}$ Unbiased? Y N

$$E\,\hat\eta_n(\underline{x}) = E\log\bar{x} \neq \log E\bar{x} = \log\lambda = \eta$$

[4]The Invariance Principle makes this very easy

Problem 3: Quasimodo has k keys; exactly one of them opens the treasure chest in Notre Dame Cathedral. When a key is requested he picks one uniformly at random, so there is a $\theta = 1/k$ probability the key will work.[5] We don't know the value of the positive integer k. Esmerelda fears that there are one hundred keys (so $\theta = 0.01$). She wants to test the hypotheses

$$H_0: \ k \geq 100 \qquad vs. \qquad H_1: \ k < 100.$$

In a fixed number $n = 4$ of tries (independently, with replacement), Esmerelda had $X = 1$ success and $n - X = 3$ failures. Not bad.

a) Which possible outcomes with four tries *are* "as extreme or more so" evidence against H_0 than hers for this alternative?

$\boxed{X \geq 1}$

Also Note that the test that rejects when $X \geq C(\alpha)$ is UMP. $C(\alpha)$ is a cutoff constant determined by the level α.

b) Find the P-value for a test of H_0 against H_1 with $X = 1$ (and $n = 4$). Do you accept or reject H_0 at level $\alpha = 0.05$?

$P =$ $.039 < .05$ ○Accept ☑Reject

$$p = P(X \geq 1 \mid k = 100) = 1 - P(X = 0 \mid k = 100)$$
$$= 1 - \left(1 - \frac{1}{100}\right)^4 \approx .039 < .05$$

You may also derive from the definition of p-values

$$p = \min\{\alpha : \text{Reject when } X = 1\} = \min\{\alpha : C(\alpha) \leq 1\}$$
$$= \min\{\alpha : P(X \geq 1 \mid k = 100) \leq \alpha\} = P(X \geq 1 \mid k = 100)$$

c) Captain Phoebus felt that the values $k = 10$ and $k = 100$ were equally likely (and that no other k was possible), before he learned of Esmerelda's experiment. Find the posterior probability:

$$\pi(k = 100 \mid X = 1) = \frac{P(X=1 \mid k=100) \cdot \pi(k=100)}{P(X=1 \mid k=100) \cdot \pi(k=100) + P(X=1 \mid k=10) \pi(k=10)}$$
$$= \frac{\binom{5}{1} \cdot \left(\frac{1}{100}\right)^1 \cdot \left(\frac{99}{100}\right)^4}{\binom{5}{1}\left(\frac{1}{100}\right)^1\left(\frac{99}{100}\right)^4 + \binom{5}{1}\left(\frac{1}{10}\right)^1\left(\frac{9}{10}\right)^4} \approx .13.$$

[5] Some may prefer to express all the dist'ns and hypotheses in terms of θ instead of k.

Problem 4: The count X comes from the Poisson distribution with pmf

$$p_\theta(k) = \mathsf{P}_\theta[X = k] = \frac{\theta^k}{k!}e^{-\theta}, \qquad k = 0, 1, 2, \ldots$$

with uncertain parameter $\theta > 0$. We wish to test the hypothesis H_0 that $\theta = 5.0$ against various alternatives. Here are the values of the pmf $p_5(k) \equiv \mathsf{P}[X = k \mid \theta = 5]$ and CDF $P_5(k) \equiv \mathsf{P}[X \le k \mid \theta = 5]$ for $0 \le k \le 10$:

k:	0	1	2	3	4	5	6	7	8	9	10
$p_5(k)$:	0.007	0.034	0.084	0.140	0.175	0.175	0.146	0.104	0.065	0.036	0.018
$P_5(k)$:	0.007	0.040	0.125	0.265	0.440	0.616	0.762	0.867	0.932	0.968	0.986

a) Consider the test of the hypotheses

$$H_0: \ \theta = 5.0 \qquad vs. \qquad H_1: \ \theta \ne 5.0$$

based on $n = 1$ observation that rejects H_0 when $X \in \mathcal{R} = \{0, 1; \quad 7, 8, 9, \ldots\}$, *i.e.*, when $X \notin \{2, 3, 4, 5, 6\}$. Find the size α of this test. Show your work (*e.g.*, tell what number(s) you used from the table above):

$\alpha =$

$1 - (.084 + .14 + .175 + .175 + .146)$ or

$= 1 - (.762 - .040) = 1 - .722 = .278$

b) Give an expression (sum or integral) for the *power* of the test described in a) above, as a function of θ:

$\mathsf{pow}(\theta) =$ $1 - e^{-\theta} \cdot \left(\frac{\theta^2}{2} + \frac{\theta^3}{6} + \frac{\theta^4}{24} + \frac{\theta^5}{120} + \frac{\theta^6}{720} \right)$

c) Find the P-value for a test of

$$H_0: \ \theta = 5.0 \qquad vs. \qquad H_1: \ \theta > 5.0$$

for a single observation of $X = 9$. Do we Accept or Reject at level $\alpha = 0.04$?

$P =$ $P(X \ge 9 \mid \theta = 5.0)$ ☑ Accept ○ Reject

$= 1 - P(X \le 8 \mid \theta = 5.0) = 1 - .932 = .068 > .04$

Problem 5: For 2.5pt each, circle "T" for True or "F" for False. No explanations are necessary.

a) T F — The Beta prior distribution is conjugate for the Geometric sampling distribution.

b) T F — If X has a χ^2_ν distribution then $\mathsf{Var}X = \nu$.

c) T F — The estimator with the smallest Mean Square Error is always unbiased.

d) T F — If $\{X_i\} \sim \mathsf{Be}(\alpha, 2)$ then $\bar{X}$ is sufficient for α.

e) T F — UMP tests exist for all one-sided hypotheses.

f) T F — Reject H_0 if you observe data X in the critical region $\mathcal{R}$.

g) T F — The *power* function is the probability that H_0 is rejected.

h) T F — If the P-value is above 0.99, *reject* at level $\alpha = 0.01$.

Problem 6: For 2.5pt each, circle "T" for True or "F" for False, or write short answers in the boxes. All questions all concern a random sample $\mathbf{x} = \{X_i\}_{1 \le i \le N}$ from the $\mathsf{No}(\mu, \sigma^2)$ distribution. Unless told otherwise, μ and σ^2 are unknown. No explanations are necessary.

a) What prior distribution is conjugate for μ, if σ^2 is known?? $\boxed{N(\mu_0, \sigma_0^2)}$

b) For fixed sample size n, a likelihood ratio test of $H_0: \ \mu = 0$ *vs.* $H_0: \ \mu > 0$ of smaller size α will have higher power $1-\beta$. T F

c) The t test of $H_0: \ \mu = 0$ *vs.* $H_1: \ \mu > 0$ is UMP. T F

~~d)~~ If $\sigma^2 = 32$, $N = 4$, $\bar{X} = 3$ with improper prior $\pi(\mu) = 1$ then what is the posterior distribution $\pi(\mu \mid \mathbf{x})$ for μ? $\boxed{N(\bar{x}, \sigma^2/N)} = N(3, 8)$

(Hard)

~~e)~~ If H_0 is true, the P-value for the t test of $H_0: \ \mu = 0$ *vs.* $H_1: \ \mu > 0$ has the $\mathsf{Un}(0,1)$ dist'n. T F

$p(\underline{x}) = P(T(\underline{x}) > t(\underline{x}) \mid H_0) = 1 - F_{t_{n-1}}(t)$

$P(p(\underline{x}) \le u) = P(1 - F_{t_{n-1}}(T) \le u) = P(F_{t_{n-1}}(T) \ge 1-u) = 1 - F_{t_{n-1}} \circ F_{t_{n-1}}^{-1}(1-u) = 1 - 1 + u = u.$

f) Use a χ^2 dist'n to test $H_0: \ \sigma^2 = 1$ against $H_1: \ \sigma^2 > 1$. T F

g) What's the dist'n of $\sum_{1 \le i \le N} (X_i - \bar{X})^2$ if $N = 9$ and $\sigma^2 = 2$? $\boxed{2\chi^2_8}$

h) Use a t distribution to estimate μ if $\sigma^2 = 1$ and $N \le 10$. T F

Problem 7: Ada and Van Veen model the numbers of potholes on different segments of Rte I-85 as independent Poisson random variables with means λL_i, where L_i denotes the length (in miles) of the segment. Thus,

$$\{X_i\} \overset{\text{ind}}{\sim} \mathsf{Po}(\lambda L_i).$$

The data are independent, but not (quite) identically distributed, because the (known, non-random) lengths L_i vary. Ada and Van observe 10 pothole counts $\mathbf{x} = \{X_i\}$ on ten road segments of specified lengths $\{L_i\}$, with empirical pothole rates X_i/L_i, as follows:

	Data										Sum
Potholes	105	72	105	85	115	110	95	100	105	108	1000
Length (mi)	25	8	30	20	23	20	19	30	15	10	200
Rate (Ph/mi)	4.2	9.0	3.5	4.25	5.0	5.5	5.0	3.33	7.0	10.8	57.58

a) Find the likelihood function for λ for the data above. Simplify!

$f(\mathbf{x} \mid \lambda) =$ $C \cdot \lambda^{1000} e^{-200\lambda}$

$$f(\underline{x}|\lambda) = \prod_{i=1}^{n} f(x_i|\lambda_i) = \prod_{i=1}^{n}\left[\frac{(\lambda L_i)^{x_i}}{x_i!} e^{-\lambda L_i}\right]$$

$$= \lambda^{\sum_i x_i} \cdot e^{-\lambda \sum_i L_i} \cdot \prod_{i=1}^{n} \frac{L_i^{x_i}}{x_i!}$$

$$= C \cdot \lambda^{1000} e^{-200\lambda}$$

b) Find the MLE for these data:

$\hat{\lambda} =$ 5

$$\log L(\lambda) = \text{Const} + 1000 \log \lambda - 200\lambda$$

$$\frac{d}{d\lambda} \log L(\lambda) = \frac{1000}{\lambda} - 200 = 0 \Rightarrow \hat{\lambda} = 5$$

$$\frac{d^2}{d\lambda^2} \log L(\lambda) = -\frac{1000}{\lambda^2} < 0. \quad \text{So } \hat{\lambda} = 5 \text{ is MLE}.$$

Problem 7 (cont):

c) Ada and Van choose a gamma prior distribution

$$\pi(\lambda) \sim \mathsf{Ga}(\alpha = 20,\ \beta = 5)$$

for the rate parameter λ. Find their posterior distribution (give its name and any parameter(s)) and their posterior mean, for these data:

$\pi(\lambda \mid \mathbf{x}) = \dfrac{205^{1020}}{\Gamma(1020)} \lambda^{1019} e^{-205\lambda}, \ (\lambda > 0).$ $\quad \mathsf{E}[\lambda \mid \mathbf{x}] = \dfrac{1020}{205} \approx 4.98$

$$\pi(\lambda \mid \underline{x}) \propto f(\underline{x} \mid \lambda) \cdot \pi(\lambda) \propto \cancel{\lambda^{1000} e^{-200\lambda}}\ \lambda^{\sum_i x_i} \cdot e^{-\lambda \sum_i L_i} \cdot \frac{\cancel{\beta}\beta^{\alpha}}{\Gamma(\alpha)} \cdot \lambda^{\alpha-1} e^{-\beta\lambda}$$

$$\propto \lambda^{(\sum_i x_i)+\alpha-1} e^{-\lambda(\beta + \sum_i L_i)}$$

$$\sim \text{Gamma}\left(\alpha + \sum_i X_i,\ \beta + \sum_i L_i\right)$$

$$= \text{Gamma}(1020,\ 205)$$

d) Briefly, *why* should $\hat{\lambda}(\mathbf{x})$ or $\mathsf{E}[\lambda \mid \mathbf{x}]$ be better estimates of the pothole rate than the average value ($\bar{R} = 5.758$) of the empirical pothole rates $R_i \equiv X_i/L_i$ given in the bottom row of the data chart?

~~Because some of the R_i's are~~

This is because the R_i's are estimated from observations corresponding to different L_i's. Intuitively, one should "weight" the observations w/ longer L_i more.

Problem 8:

In a study of drug treatment for depression, subjects are given a "mood test" twice— once before, and once after, treatment with a drug (either Prozac or a placebo). Scores range from zero to twenty, with higher scores indicating a cheerier mood and less depression. We will assume scores are approximately normally distributed, with the same variance. The scores for the nine subjects in the treated (Prozac) group, along with the sample means $\bar{x}$ and sums-of-squares SSQ $= \sum(x_i - \bar{x})^2$, $\sum(y_i - \bar{y})^2$ were:

Subject: $i =$	1	2	3	4	5	6	7	8	9	Avg	SSQ
Pre-med X_i	3	0	6	7	4	3	2	1	4	3.333	40
Post-med Y_i	5	1	5	7	10	9	7	11	8	7.000	74
Change d_i	2	1	−1	0	6	6	5	10	4	3.667	98

a) To test the hypothesis of "no change" against the alternative that Prozac *improves* mood, what would be the *alternate* hypothesis?

$H_0 : \mu_x = \mu_y$ $\qquad$ $H_1 :$ $\mu_x < \mu_y$

b) Which test would you recommend, and *why*?

☑ Paired t ○ χ^2 ○ Two-sample t ○ F ○ Normal

c) How many degrees of freedom does the test you recommend have?

$\nu =$ $n-1 = 8$.

Problem 8 (cont):

As before, the mood test data are:

Subject: $i =$	1	2	3	4	5	6	7	8	9	Avg	SSQ
Pre-med X_i	3	0	6	7	4	3	2	1	4	3.333	40
Post-med Y_i	5	1	5	7	10	9	7	11	8	7.000	74
Change d_i	2	1	-1	0	6	6	5	10	4	3.667	98

d) Find a 90% confidence interval for the change in mood score, $\mu_y - \mu_x$:

$$\cancel{T = \frac{\bar{d} - (\mu_Y - \mu_X)}{SSQ_d / \sqrt{n}}} \qquad T = \frac{\bar{d} - (\mu_Y - \mu_X)}{S_d / \sqrt{n}} = \frac{\bar{d} - (\mu_Y - \mu_X)}{\sqrt{\frac{SSQ_d}{(n-1) \cdot n}}}$$

$$.9 = P\left(F_{t_{n-1}}(.05) \le \frac{\bar{d} - (\mu_Y - \mu_X)}{\sqrt{\frac{SSQ_d}{n(n-1)}}} \le F_{t_{n-1}}(.95) \right)$$

$$= P\left(\bar{d} - F_{t_{n-1}}(.95) \sqrt{\frac{SSQ_d}{n(n-1)}} \le \mu_Y - \mu_X \le \bar{d} + F_{t_{n-1}}(.95) \sqrt{\frac{SSQ_d}{n(n-1)}} \right)$$

So a 90% CI for $\mu_Y - \mu_X$ is

$$\left[\bar{d} - F_{t_{n-1}}(.95) \cdot \sqrt{\frac{SSQ_d}{n(n-1)}} \; , \; \bar{d} + F_{t_{n-1}}(.95) \cdot \sqrt{\frac{SSQ_d}{n(n-1)}} \right]$$

$$\left[3.667 - 1.860 \times \sqrt{\frac{98}{9 \times 8}} \; , \; 3.667 + 1.860 \times \sqrt{\frac{98}{9 \times 8}} \right] = [1.497, 5.837]$$

e) (1pt) Any concerns about the assumptions of normality and equal variance?

① Scores are bounded below by 0. ② discreteness in the scores.

Problem 9: Holden Caulfield and Susie Creamcheese wonder whether the color of jelly beans is related to their heat tolerance. They decide to take a scientific approach: they put each jelly bean into a micro-wave oven for 10 seconds, and recorded what happened. Here are their data:

	Blue	White	Black	total
Nothing	20 (15)	4 (10)	6 (5)	30
Melted	8 (12)	12 (8)	4 (4)	24
Flames	2 (3)	4 (2)	0 (1)	6
total:	30	20	10	60

group group

a) Which test would you recommend?

○ Paired t ✓ χ^2 ○ Two-sample t ○ F ○ Normal

b) How many degrees of freedom does the test you recommend have?

$\nu =$ $9-1-(3-1)-(3-1)-2 = 2$ d.f. or $(3-1)\times(3-1)-2 = 2$. ↑ merged cells.

c) Perform the test— *i.e.*, specify the null and alternate hypotheses, pick a test statistic, find its value and its approximate dist'n under H_0, *etc.* Show your work. Do you accept or reject at level $\alpha = 0.05$? ○ Acc ○ Rej

$$\chi^2 = \frac{(20-15)^2}{15} + \frac{(4-10)^2}{10} + \frac{(6-5)^2}{5} + \frac{(8-12)^2}{12} + \frac{(12-8)^2}{8} + \frac{(4-5)^2}{5} + \frac{(6-5)^2}{5}$$

$$\approx 9.2.$$

χ^2_2 dist^n is also the Gamma$(1, \frac{1}{2})$ dist^n so its pdf is $f(x) = \frac{\frac{1}{2}}{1} \cdot x^0 \cdot e^{-\frac{1}{2}x} = \frac{1}{2}e^{-\frac{1}{2}x}$

d) (XC): Give the P-value to six correct digits. Show your work.

So $F(x) = 1 - e^{-\frac{1}{2}x}$

Thus p-value $= 1 - F(9.2) = e^{-4.6} \approx .0100518$

Problem 10: The Pareto distribution is often used to model "heavy-tailed" data like incomes or storm intensities where some observations are *much* larger than others. Appropriately scaled, it has pmf and CDF

$$f(x) = \theta x^{-\theta-1}\mathbf{1}_{\{x\ge 1\}} \qquad F(x) = \begin{cases} 0 & x \le 1 \\ 1 - x^{-\theta} & x > 1 \end{cases}$$

a) Find the Maximum Likelihood Estimator for a sample $\mathbf{x}$ of size n:
$\hat{\theta}_n(\mathbf{x}) = \frac{n}{\sum_{i=1}^{n} \log x_i}$

$L(\theta) = \theta^n \left(\prod_{i=1}^{n} x_i\right)^{-\theta-1} \cdot \mathbb{1}(\min x_i \ge 1)$.

$\log L(\theta) = n \log\theta - (\theta+1)\sum_{i=1}^{n} \log x_i$

$\frac{d \log L(\theta)}{d\theta} = \frac{n}{\theta} - \sum_{i=1}^{n} \log x_i = 0$

$\Rightarrow \hat{\theta} = \frac{n}{\sum_{i=1}^{n} \log x_i}$. $\Bigg| \frac{d^2 \log L(\theta)}{d\theta} = -\frac{n}{\theta^2} < 0$

So it's MLE

~~b)~~ Is this an *exponential family*? If so, give the natural sufficient statistic $T_n(\mathbf{x})$ for a sample $\mathbf{x}$ of size n; if not, why not?
$T_n(\mathbf{x}) = -\log x$. Expo Fam? Y N

$f(x|\theta) = \theta x^{-\theta-1}\mathbb{1}(x\ge 1) = e^{-(\theta+1)\log x + \log\theta} \cdot \mathbb{1}(x \ge 1)$

$= e^{-\theta \log x + \log\theta} \cdot \frac{\mathbb{1}(x\ge 1)}{x}$

c) Find the P-value for the Likelihood Ratio Test of

$$H_0: \theta = 2 \ (\theta_0) \quad vs. \quad H_1: \theta = 1 \ (\theta_1)$$

for the single observation of $x = 20$:
$P = \frac{1}{400}.$

The LR test rejects when

$$\frac{f(x|\theta_1)}{f(x|\theta_0)} = \frac{x^{-2} \cdot \mathbb{1}(x\ge 1)}{2 \cdot x^{-3}\mathbb{1}(x \ge 1)} = \frac{x}{2} \ (\text{for } x \ge 1) > k.$$

So it rejects when $X > c$.

$$\text{p-val} = P(X \ge 20 | \theta_0) = 1 - P(X \le 20 | \theta_0) = 1 - (1 - 20^{-2}) = \frac{1}{400}.$$

Problem 10 (cont):

~~d)~~ Find the posterior mean of θ with an improper uniform prior distribution $\pi(\theta) \equiv \mathbf{1}_{\{\theta>0\}}$ for a sample of size $n = 1$ of $x = 20$ (hint: $a^b = e^{b \log a}$). Simplify; no integration is needed.

$\mathsf{E}[\theta \mid X = 20] = \frac{1}{\log 20}$

For $x = 20$.

$$\pi(\theta \mid x) \propto \pi(\theta) \cdot f(x \mid \theta)$$

$$= 1 \cdot \theta x^{-\theta-1}$$

$$\propto \theta e^{-\theta \log x} \quad \text{for } \theta > 0.$$

$$\sim Exp(\log x) = Exp(\log 20).$$

$$E(\theta \mid X = 20) = \frac{1}{\log 20}$$

e) Show that the $\mathsf{Ga}(\alpha, \beta)$ family are conjugate prior distributions for θ. Give the parameter values for the posterior, and the *posterior mean* for a sample of size n (using your sufficient statistic T_n from part b) above):

$\pi(\theta \mid \mathbf{x}) \sim \mathsf{Ga}(\alpha^*, \beta^*)$ with:

$\alpha^* = \alpha + 1$ $\quad \beta^* = \beta + \log x$ $\quad \mathsf{E}[\theta \mid \mathbf{x}] = \frac{\alpha^*}{\beta^*} = \frac{\alpha+1}{\beta + \log x}$,

$$\pi(\theta \mid x) \propto \pi(\theta) \cdot f(x \mid \theta)$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \theta^{\alpha-1} e^{-\beta\theta} \cdot \theta x^{-\theta-1}$$

$$\propto \theta^{(\alpha+1)-1} \cdot e^{-(\beta + \log x)\theta}$$

$$\sim Gamma(\alpha+1, \beta + \log x).$$

Done! Have a great summer.