MTH 375
Winter 2022
Final Exam Key

1. Suppose the random variables $X_1, \ldots, X_n$ are iid, with common density

$$f_X(x) = \begin{cases} \theta(\theta+1)x(1-x)^{\theta-1} & \text{if } 0 < x < 1, \\ 0 & \text{otherwise,} \end{cases}$$

where θ is an unknown parameter. Find a sufficient statistic for θ.

$$L(\theta) = f_{X_1,\ldots,x_n}(x_1, \ldots, x_n) = \left[\theta^n(\theta+1)^n \left(\prod_{i-1}^{n}(1-x_i)\right)^{\theta-1}\right] \cdot \left[\prod_{i=1}^{n} x_i\right].$$

The first factor is $g(T, \theta)$. The sufficient statistic is $T = \prod_{i=1}^{n}(1-x_i)$.

2. Let X be geometric with parameter p. You are testing the hypotheses

$$H_0 : p = .8 \text{ vs. } H_a : p < .8.$$

and decide that the test will be "Reject H_0 if $X > 2$."

(a) What is the level of significance of this test?

Under H_0, $\alpha = P(X > 2) = \sum_{n=3}^{\infty} .8(.2)^{n-1} = \frac{.8\cdot(.2)^2}{1-.2} = .04$.

(b) Find a simple formula for the power function.

$\beta(p) = P(X > 2) = \sum_{n=3}^{\infty} p(1-p)^{n-1} = \frac{p\cdot(1-p)^2}{1-(1-p)} = (1-p)^2$.

3. Suppose $Y_1, \ldots, Y_{10}$ are iid with common pdf

$$f(y) = \begin{cases} 2y/\theta^2 & \text{if } 0 < y < \theta \\ 0 & \text{otherwise} \end{cases}$$

for some $\theta > 0$.

(a) Show that $Q = \dfrac{\max\{Y_1, \cdots, Y_{10}\}}{\theta}$ is a pivotal quantity.

$$F_Q(t) = P\left(\frac{\max\{Y_1, \cdots, Y_{10}\}}{\theta} \le t\right) = P\left(\max\{Y_1, \cdots, Y_{10}\} \le \theta t\right) = (P(Y_k \le \theta t))^{10}$$

$= \left(\int_0^{\theta t} \frac{2y}{\theta^2}\, dt\right)^{10} = \left(\frac{(\theta t)^2}{\theta^2}\right)^{10} = t^{20}$, which does not depend on θ.

(b) Find a 90% confidence interval for θ based on Q.

$$.90 = P\left(.05^{1/20} \le \frac{Y_{(10)}}{\theta} \le .95^{1/20}\right) = P\left(\frac{1}{.95^{1/20}} \le \frac{\theta}{Y_{(10)}} \le \frac{1}{.05^{1/20}}\right)$$

$$= P\left(\frac{Y_{(10)}}{.95^{1/20}} \le \theta \le \frac{Y_{(10)}}{.05^{1/20}}\right) \approx P\left(\frac{Y_{(10)}}{.997} \le \theta \le \frac{Y_{(10)}}{.861}\right).$$

Either of these last inequalities define the required confidence interval.

4. A coin with probability p of landing 'heads' is tossed 5 times. A Bayesian statistician believes that p may be 1/4, 1/2 or 3/4, and assigns prior probabilities of $\pi(1/4) = 1/6$, $\pi(1/2) = 1/3$ and $\pi(3/4) = 1/2$.

The coin is tossed 5 times, and the result is 3 heads and 2 tails.

(a) What is the posterior distribution of p?

Let X be the number of heads tossed. Then

$$P(p = 1/4 | \, X = 3) = \frac{\pi(1/4)P(X = 3 \mid p = 1/4)}{P(X = 3)} = \frac{(1/6) \cdot \binom{5}{3}(1/4)^3(3/4)^2}{P(X = 3)} \approx \frac{.015}{P(X = 3)},$$

$$P(p = 1/2 | \, X = 3) = \frac{\pi(1/2)P(X = 3 \mid p = 1/2)}{P(X = 3)} = \frac{(1/3) \cdot \binom{5}{3}(1/2)^3(1/2)^2}{P(X = 3)} \approx \frac{.104}{P(X = 3)},$$

$$P(p = 3/4 | \, X = 3) = \frac{\pi(3/4)P(X = 3 \mid p = 3/4)}{P(X = 3)} = \frac{(1/2) \cdot \binom{5}{3}(3/4)^3(1/4)^2}{P(X = 3)} \approx \frac{.132}{P(X = 3)}.$$

$P(X = 3)$ is the sum of the numerators, so we have

$$P(p = 1/4 \mid X = 3) \approx \frac{.015}{.251} \approx .058,$$

$$P(p = 1/2 \mid X = 3) \approx \frac{.104}{.251} \approx .416,$$

$$P(p = 3/4 \mid X = 3) \approx \frac{.132}{.251} \approx .526.$$

This is the posterior distribution of p.

(b) Assuming a quadratic loss function (i.e., $L(\hat{p}, p) = (\hat{p} - p)^2$), what is the Bayesian estimate of p?

The Bayesian estimate of p is $E(p \mid X = 3) = 1/4 \cdot 058 + 1/2 \cdot .416 + 3/4 \cdot .526 \approx .617$. (The exact answser is 95/154.)

5. Suppose the pairs $(X_1, Y_1), \ldots, (X_{10}, Y_{10})$ satisfy the usual hypotheses for linear regression, i.e., $Y = \beta_0 + \beta_1 X + \varepsilon$ and so on. Ten such points turn out to be

X_i	.01	.03	.03	.19	.24	.25	.66	.77	.86	.94
Y_i	.87	.31	.60	.24	.67	.16	.80	.95	.69	.64

(a) Test at level of significance .05 the hypotheses

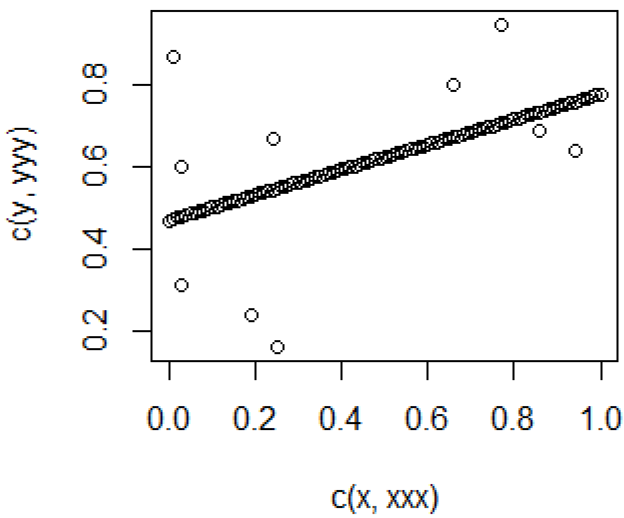
$$\mathrm{H}_0 : \beta = 0 \quad \text{vs.} \quad \mathrm{H}_1 : \beta \neq 0.$$

```
Coefficients:
            Estimate   Std. Error t value Pr(>|t|)
(Intercept)   0.4707     0.1245   3.780   0.00539
x             0.3074     0.2349   1.309   0.22701

> qt(df=8,c(.025,.975))
[1] -2.306004  2.306004
```

The test is "Reject $H_0 : \beta_1 = 0$ if $|t_{\hat{\beta}_1}| > 2.31$". According to R, the t-value of $\hat{\beta}_1$ is 1.309, so we do not reject H_0. In words: The data does not show (at level of significance $\alpha = .05$) that X, Y are linearly related.

(b) Plot the points and the regression line, and explain why the result of (a) is reasonable.



The test result is plausible because, while there seems to be a general tendency for y to rise as x rises, the data is more scattered than adhering to the regression line. The line may very well be an accident.

6. Let $X_1, \cdots, X_n$ be iid normal with unknown μ and $\sigma = 1$.

(a) Find an unbiased estimator for e^μ.

The MGF of a normal(μ, σ^2) random variable Y is $M_Y(t) = E(e^{tY}) = e^{\mu t + \frac{1}{2}\sigma^2 t}$.

We know that $\overline{X}$ is normal$(\mu, 1/n)$, its MGF is $M_{\overline{X}}(t) = E(e^{t\overline{X}}) = e^{\mu t + \frac{t}{2n}}$.

Let $t = 1$; we have $E(e^{\overline{X}}) = e^{\mu + \frac{1}{2n}}$.

Finally, solve for e^μ to obtain $E(e^{\overline{X} - \frac{1}{2n}}) = e^\mu$.

The required unbiased estimator is $T = e^{\overline{X} - \frac{1}{2n}}$.

(b) Is your estimator an UMVUE for e^μ? Why or why not?

Yes; T is an unbiased function of a sufficient statistic ($\overline{X}$), so T is an UMVUE.