Solutions to Final

1. (a) The t statistic for this one sided test is

$$t = (\text{slope estimate})/(\text{slope standard error}) = -0.006328/0.001538 = -4.114434$$

which has a t distribution with 4 degrees of freedom if the true slope is zero. The one sided test rejects if this value is small, so the p-value is the probability that a t random variable with 4 degrees of freedom is less than or equal to -4.114434. From the tables this value is between 0.005 and 0.01. Alternatively, this value is half the two-sided p-value reported in the output, so the p-value is $0.01467/2 = 0.007335$.

(b) The 97.5th percentile of the t_4 distribution is 2.776, so a 95% confidence interval for the slope is

$$\begin{aligned}(\text{slope estimate}) \pm (\text{slope standard error}) \times 2.776 &= -0.006328 \pm 0.001538 \times 2.776\\ &= -0.006328 \pm 0.004269\\ &= [-0.0106, -0.00206]\end{aligned}$$

2. The ratio of the likelihoods for $0 \leq x \leq 1$ is $f_1(x)/f_0(x) = 2(1-x)$. The most powerful test rejects when this ratio is large, which is equivalent to rejecting when X is small. Rejecting when $X \leq 0.2$ produces a Type I error probability of $\alpha(\delta) = 0.2$. So the critical region $C = \{x : x \leq 0.2\}$ minimizes $\beta(\delta)$ among all tests of size at most 0.2. The Type II error probability for this test is

$$\begin{aligned}\beta(\delta) = \Pr(X > 0.2|H_1) &= \int_{0.2}^{1} f_1(x)dx\\ &= \int_{0.2}^{1} 2(1-x)dx = \left[-(1-x)^2\right]_{0.2}^{1}\\ &= (0.8)^2 = 0.64\end{aligned}$$

3. The expected counts assuming homogeneity are

	Faculty	Librarians	Total
Male	56.45455	58.54545	115
Female	51.54545	53.45455	105
Total	108	112	220

The χ^2 statistic is thus

$$\begin{aligned}Q &= \frac{(56-56.45455)^2}{56.45455} + \frac{(59-58.54545)^2}{58.54545} + \frac{(52-51.54545)^2}{51.54545} + \frac{(53-53.45455)^2}{53.45455}\\ &= 0.01506\end{aligned}$$

The 95th percentile of the χ^2 distribution with one degree of freedom is 3.841. So the test does not reject at the 0.05 level (or any other reasonable level: the p-value is approximately 0.9).

4. (a) The likelihood is

$$f_n(y_1,\dots,y_n|\theta) = \prod_{i=1}^n \frac{(\theta t_i)^{y_i}}{y_i!} e^{-\theta t_i} \propto \theta^{\sum y_i} \exp\{-\theta \sum t_i\}$$

For $\sum y_i = 0$ this is decreasing and hence maximized by $\widehat{\theta} = 0$. For $\sum y_i > 0$ the likelihood is continuously differentiable, is zero at $\theta = 0$ and tends to zero as θ tends to infinity, so a global mazimizer exists in the range $(0,\infty)$ and must be a zero of the derivative of the log likelihood. The derivative of the log likelihood is

$$\theta \sum y_i - \sum t_i$$

This has a unique zero at $\widehat{\theta} = \sum y_i / \sum t_i$. Thus in both cases the maximum likelihood estimator is given by $\widehat{\theta} = \sum y_i / \sum t_i$.

(b) The expected value of $\widehat{\theta}$ is

$$E[\widehat{\theta}] = \frac{\sum E[Y_i]}{\sum t_i} = \frac{\sum \theta t_i}{\sum t_i} = \theta$$

for all θ, so $\widehat{\theta}$ is unbiased. Its variance is

$$\text{Var}(\widehat{\theta}) = \frac{\text{Var}(\sum Y_i)}{(\sum t_i)^2} = \frac{\sum Var(Y_i)}{(\sum t_i)^2} = \frac{\sum \theta t_i}{(\sum t_i)^2} = \frac{\theta}{\sum t_i}$$

(c) The Fisher information in Y_i is

$$I^{(i)}(\theta) = -E[\lambda''(Y_i|t_i,\theta)] = -E[-Y_i/\theta^2] = E[Y_i]/\theta^2 = t_i/\theta$$

So the Fisher information in the sample is

$$I(\theta) = \sum I^{(i)}(\theta) = \sum t_i/\theta = \frac{\sum t_i}{\theta}$$

Since $\text{Var}(\widehat{\theta}) = 1/I(\theta)$, the maximum likelihood estimator is efficient.

(d) For $\theta_2 > \theta_1$ the likelihood ratio is

$$\frac{\theta_2^{\sum Y_i} \exp\{-\theta_2 \sum t_i\}}{\theta_1^{\sum Y_i} \exp\{-\theta_1 \sum t_i\}} = \left(\frac{\theta_2}{\theta_1}\right)^{\sum Y_i} \exp\{(\theta_1 - \theta_2) \sum t_i\}$$

This is monotone increasing in $\sum Y_i$, so the UMP test of these hypotheses rejects the null hypothesis when $\sum Y_i$ is large.

Summary of Scores

Mean	45.21
SD	5.50
Min	29
Lower Quartile	45.5
Median	47
Upper Quartile	48.5
Max	50