Statistics 135, Fall 2004
Solutions to Final Exam
December 17, 2004

1. (a) $\bar{X}_s = W_H \bar{X}_H + W_L \bar{X}_L$, where $W_H = 1/6$, $W_L = 5/6$.

   (b) $\operatorname{Var} \bar{X}_s = W_H^2 \operatorname{Var} \bar{X}_H + W_L^2 \operatorname{Var} \bar{X}_L$, and

   $$\operatorname{Var} \bar{X}_i = \frac{\sigma_i^2}{n_i}\left(1 - \frac{n_i - 1}{N_i - 1}\right).$$

   Plugging in values gives a standard error of .677.

   (c) For proportional allocation, $n_H = 50$ and $n_L = 250$. Repeating part (b) with these numbers gives a standard error of .707, which is worse. (This is not surprising: out analysis of the Neyman optimal allocation showed that its better to oversample more variable strata.)

2. (a) $L(\theta) = \left[\binom{3}{0}\theta^0(1-\theta)^3\right]\left[(1-\theta)^3\theta\right] = \theta(1-\theta)^6$.

   (b) Differentiating the log-likelihood and setting equal to 0 gives $\hat{\theta} = 1/7$.

3. (a) $L(\theta) = (\theta+1)^n \prod_{i=1}^n x_i^\theta$, and the log-likelihood $l(\theta)$ is given by $n\log(\theta+1) + \theta\sum_i \log X_i$. Differentiating $l(\theta)$ and setting equal to 0 gives $\hat{\theta} = -(1 + n/\sum_i \log X_i)$.

   (b) We may use either $\mathbb{E}\big(l'(\theta)\big)^2$ or $-\mathbb{E}l''(\theta)$. Since the former requires finding moments for $\log X$, the latter is much easier, and gives $(\theta+1)^2/n$, which can be estimated by $(\hat{\theta}+1)^2/n = n/(\sum_i \log X_i)^2$. (To do the former, you must use change-of-variables to show $-\log X \sim \operatorname{Exp}(\theta+1)$. You then get the same result after plugging in exponential moments and cleaning up the mess.)

4. (a) Here and throughout take $H_0 : \mu = 100$ and $H_1 : \mu = 125$. The likelihood ratio is then

   $$\begin{aligned}\lambda &= \frac{p_0(120)}{p_1(120)} \\ &= \frac{(2\pi\sigma^2)^{-1/2}\exp(-(120-\mu_0)^2/2\sigma^2)}{(2\pi\sigma^2)^{-1/2}\exp(-(120-\mu_1)^2/2\sigma^2)} \\ &= \exp\{-(20^2 - 5^2)/2\sigma^2\} \\ &= .74\end{aligned}$$

   (b) Let $\theta \sim \text{Bernoulli}(1/2)$, so that $X|_\theta \sim \mathcal{N}(\mu_\theta, 25^2)$. Then

   $$\begin{aligned}\mathbb{P}(\theta = 1 | X = 120) &= \frac{\mathbb{P}(\theta=1)p(120|\theta=1)}{p(120)} \\ &= \frac{\mathbb{P}(\theta=1)p_1(120)}{\mathbb{P}(\theta=0)p_0(120) + \mathbb{P}(\theta=1)p_1(120)} \\ &= \frac{p_1(120)}{p_0(120)+p_1(120)},\end{aligned}$$

   since $\mathbb{P}(\theta = 0) = \mathbb{P}(\theta = 1)$. Evaluating the last line gives .574.

(c) $\mathbb{P}_0(X > 125) = \mathbb{P}_0\big((X - 100)/25 > (125 - 100)/25\big) = 1 - \Phi(1) = .159$. (So this test has α much higher than what we're usually comfortable with.)

(d) $\mathbb{P}_1(X > 125) = .5$ since $\mu_1 = 125$.

(e) The p-value is $\mathbb{P}_0(X > 120) = \mathbb{P}_0\big((X - 100)/25 > (120 - 100)/25\big) = 1 - \Phi(.8) = .212$.

5. (a) We will use a two-sample t test with unequal variances (although $\mu_T = 9$, $\sigma_T = 5$ suggests a skewed distribution for the treatment group — since the times cannot be negative — so we might have preferred a non-parametric method had the data been provided).

$$\begin{aligned} s_{\bar{X}_T - \bar{X}_C} &= \sqrt{\frac{s_T^2}{n_T} + \frac{s_C^2}{n_C}} \\ &= 1.08 \end{aligned}$$

(b) For $H_0 : \mu_T = \mu_C$ vs. $H_1 : \mu_T \neq \mu_C$,

$$T = \frac{\bar{X}_T - \bar{X}_C}{s_{\bar{X}_T - \bar{X}_C}} = -.93.$$

(c) The textbook suggests comparing T to $\mathcal{T}_{31}$, where the 31 comes from a rather complicated formula. You were not provided with a t-table, so there was no point in computing this number. $\mathcal{T}_k$ has heavier tails than the standard normal for all k, so looking up the p-value on a normal table provides a lower bound — and a reasonably tight bound for large n. Using the normal table gives a two-tailed p-value of $2 \times .1768 = .3536$. (The p-value from a t-table with $df = 31$ would have been .3602.) We therefore do not reject H_0.

For full credit, your answer must have included some discussion of why using a normal table was acceptable. Both one-tailed and two-tailed answers are appropriate.

6. The design matrix is

$$X = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}$$

Let $p = (p_1, p_2, p_3)^T$. The least squares estimate of p is

$$\hat{p} = (X^T X)^{-1} X^T Y$$

7. (a) $\hat{Y} = X\hat{\beta} = PY$, where $P = X(X^T X)^{-1} X^T$.
Using $P = P^T = P^2$, $\Sigma_{\hat{Y}\hat{Y}} = \sigma^2 P$

(b)

$$\begin{aligned}
\sum_{i=1}^{n} Var(\hat{Y}_i) &= tr(\Sigma_{\hat{Y}\hat{Y}}) \\
&= \sigma^2 tr(P) \\
&= \sigma^2 tr(X(X^TX)^{-1}X^T) \\
&= \sigma^2 tr(X^TX(X^TX)^{-1}) \\
&= p\sigma^2
\end{aligned}$$

8. $U = (1\ 1\ 1\ 1)Z$ and $V = (1\ 1\ -1\ -1)Z$

$$\begin{aligned}
\Sigma_{UV} &= (1\ 1\ 1\ 1)\Sigma_{ZZ}(1\ 1\ -1\ -1)^T \\
&= \sigma^2(1\ 1\ 1\ 1)(1\ 1\ -1\ -1)^T \\
&= 0
\end{aligned}$$

9. Gender difference can be tested via a chi-square test for independence. (Note that in the survey, only the grand total is fixed.) The null hypothesis states that the gender category is independent of the work attitude category. Let O_{ij} be the observed count in cell ij in the table and let $E_{ij} = n_{i.}n_{.j}/n_{..}$ be the expected count under the null hypothesis. The test statistic is

$$X^2 = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Alternatively, the test statistic

$$-2 \log \Lambda = 2 \sum_i \sum_j O_{ij} \log \frac{O_{ij}}{E_{ij}}$$

could be used. In either case, the approximate distribution of the test statistic under the null hypothesis is chi-square with (5-1)(2-1) = 4 degrees of freedom. The observed value of the test statistic would be compared to the quantiles of the chi-square distribution with 4 degrees of freedom.

10. For $b = 1, 2, \ldots, B$ choose a random sample of size 50 with replacement from the pairs (x_i, y_i). ($B = 1000$ would be sufficiently large.) For each of these B samples, compute the correlation coefficient r_b and let $\bar{r}$ be their average. The standard error of r can then be estimated by

$$SE(r) = \sqrt{\frac{1}{B}\sum_{b=1}^{B}(r_b - \bar{r})^2}$$