LING82100: final exam solution

1 Scales of measurement

- GRE score is *ordinal.* I believe the test is adaptive, so the score is not exactly the count of correct answers, but I will also accept *interval*; note that zero is not meaningful (the lowest score is 130).

- Word frequency is *ratio.*

- Teaspoons of baking powder is *ratio.*

- Species of *Iris* is *categorical.*

- City gas mileage is *ratio.*

- Likert scale ratings are *ordinal*, though sometimes they are treated as *interval* data as well.

2 Study design

- This is a *true experiment.*

- Because subjects hear both inflected and uninflected words, the primary manipulation—inflectional status—is *within-subjects.*

- This study calls for random intercepts for subject and item, and possibly also per-subject random slopes for inflectional status and trial number.

3 Standard error and confidence intervals

The standard error is 0.17, and the 95% confidence interval is $[21.77, 22.43]$.

4 Power analysis

A and *B* have the same probability of type I error: α. *B* has a lower probability of type II error than *A* (in fact, the power of an experiment is one minus the probability of type II error, and vice versa).

5 Test interpretation

5.1 The binomial test

A binomial test of null hypothesis that the probability that a man would drink Jeppson's Malört is 1/50 was significant ($n = 8$, $x = 2$, $p = .01$, 95% confidence interval $[.03, .65]$). We therefore reject the null hypothesis at $\alpha = .05$.

5.2 The *t*-test

```
> t.test(length ~ regularity, data = d)

        Welch Two Sample t-test

data:  length by
t = 1.8896, df = 270.56, p-value = 0.05988
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.006751323  0.329168463
sample estimates:
mean in group irregular    mean in group regular
            4.534591                 4.373383
```

An unequal variance two-sample *t*-test (or "Welch's *t*-test") was non-significant at $\alpha = .05$ ($t = 1.89$, d.f. $= 270.56$, $p = .06$). We therefore fail to reject the null hypothesis.

5.3 Logistic regression

Test results are shown in Table 1. There were significant effects, at $\alpha = .05$, of social class and emphasis. A Tukey post-hoc test on social class found a significantly higher rate of [str] in the upper middle class than the other two classes: $[\text{WC} = \text{LMC}] < \text{UMC}$.

5.4 Mixed effects linear regression

Test results are shown in Table 2. There were significant main effects, at $\alpha = .05$, of both phoneme (/t/ has a longer VOT) and language (English has a longer VOT).

	Coef.	SE	χ^2	$p(\chi^2)$
(Intercept)	1.498	0.20		
Class:			23.58	$< .001$
WC	−0.891			
LMC	−0.371	0.24		
UMC	1.262	0.32		
Emphasis:			22.33	$< .001$
less	−0.796	0.18		
more	0.796			

Table 1: Results for subsection 5.3.

	Coef.	SE	χ^2	$p(\chi^2)$
(Intercept)	4.706	3.13		
Language:			26.94	$< .001$
English	27.778	3.13		
Spanish	−27.778			
Phoneme:			26.39	$< .001$
/t/	26.927			
/d/	−26.927	3.12		

Table 2: Results for subsection 5.4.

.