For final exam, I will bring the statistical tables we have used in class.

Possible topics covered:

1. One way ANOVA
   - hypotheses and assumptions
   - construction of ANOVA table - sums of squares, degrees of freedom, additivity relationships for sums of squares and degrees of freedom; mean squares
   - observed test statistic (F_{obs}) and p-value
2. Two way ANOVA
   - hypotheses and assumptions
   - construction of ANOVA table - degrees of freedom, additivity relationships for SS and df
   - what is interaction? test for interaction first, then decide if need to test for main effects.
3. Multiple regression
   - multiple regression model, assumptions.

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + ... + e$$

   - interpreting the estimated model - e.g. predicting the mean; the β coefficient of an explanatory variable x is the change in the dependent variable y for a unit increase in this x while holding all other explanatory variables constant.
   - hypothesis tests for individual coefficients (t test). to test the hypothesis $H_0 : \beta_j = 0$ against the two sided alternative, calculate the observed value of the test statistic

$$t = \frac{\hat{\beta}_j}{s.e.(\hat{\beta}_j)}$$

   - the p-value is $2P(t_{n-1-q} > |t_{obs}|)$.
   - the degrees of freedom are $n - 1 - q$ and q is the number of the explanatory variables in the model.
   - confidence intervals for individual coefficients. A $100(1-\alpha)\%$ CI for β_j is $\hat{\beta}_j \pm t_{\alpha/2,n-1-q} s.e.(\hat{\beta}_j)$
   - overall F test. what is really being tested?
   - coefficient of determination (R^2). what is it; how is it calculated.
4. Logistic regression
   - logistic regression model:

$$log\left(\frac{p}{1-p}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + ...$$

   - interpreting coefficients. β is log odds ratio associated with a unit increase in x. how to recover probability p from log odds.

- confidence intervals and tests for β.
- manipulating to get from β to odds ratio.

5. Survival analysis

- survivor function $S(t)$
- Kaplan-Meier (product limit) estimate
- Cochran-Mantel-Haenszel (log rank) test for equality of two survivor functions: hypothesis tested, calculation of test statistic.
- Cox proportional hazards model:

$$h(t) = h_0(t)e^{b_1x_1+b_2x_2+b_3x_3+...}$$

  where $h_0(t)$ is the baseline hazard, $x_1, x_2, x_3, ...$ are covariates.
- hazard, hazard ratio, interpretation of coefficients.

6. Meta analysis

- fixed effects model only. calculation of weighted estimate, with associated confidence interval.

$$\hat{\theta} = \frac{\sum W_i Y_i}{\sum W_i}$$

  where $W_i = 1/s_i^2$, with s_i^2 the sample variance from the i'th study. The variance of θ is $1/\sum_i W_i$. The confidence interval for the weighted estimate is

$$\hat{\theta} \pm Z_{\alpha/2}\sqrt{1/\sum_i W_i}$$

7. Equivalence, non-inferiority, superiority tests

- hypotheses
- test using confidence intervals (remember to double confidence level for non-inferiority and superiority tests)