I. Multivariate Methods: Correlation and Regression

1. Same idea: Draw new observations from old multivariate distribution.
   a. Calculate test statistic whose distribution is desired for each new sample,
   b. Use methods from earlier to get confidence interval for expectation of this statistic
      i. Percentile
      ii. BCa
      iii. etc.
   c. Use for correlation, or regression \( Y_j = \beta_0 + \beta_1 X_j + \epsilon_j \), \( \epsilon_j \) iid.
   d. More generally, \( Y = X \beta + \epsilon \), with \( X = \begin{pmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_n \end{pmatrix} \).

2. New idea: Keep old X’s, but draw new residuals.
   a. Why? Maybe covariates were fixed by design
   b. How?
      i. Fit regression line
      ii. Calculate residuals and fitted values
      iii. Resample from residuals
   iv. Remember from 960:563 that residuals are not IID?
      • Sum to 0
      • Residuals associated with extreme values of explanatory variables are more variable than others.
   v. Partial Solution:
      i. Same assumption as for MWW test.
      ii. Calculate residuals and fitted values
      iii. Fitted values are group means.
      iv. Residuals are differences from group means.
      v. Estimate of shift parameter is difference in group averages.
      vi. Standard error of shift estimate is the usual pooled estimate of standard deviation, times the square root of the sum of the sample sizes.
   d. Writing as \( t \)-test as below is generally faster.

2. Special NonCase: Two Sample Testing with Unequal Error Distribution
   a. Not a special case of the regression model, since variances are unequal
   b. Bootstrap approach is pretty general: don’t assume that any aspect of distribution in two groups is the same.
   c. Proceed as above, except resampling of residuals is done within group.
      i. So, if group sizes are \( m \) and \( n \), the total number of possible bootstrap samples is \( m^m n^n \) rather than \( (m + n)^{m+n} \).
      ii. You can resample Ys instead of residuals, since items resampled in each group have the same mean.
   3. \( K \)-sample testing, for \( K > 2 \)
      a. Same as earlier:
         i. Calculate residuals
         ii. Resample new residuals from original residuals
         iii. Add residuals to fitted values to get new data
      iv. Proceed with analysis.
   b. Sample tasks:
      i. Can do inference on separate pre-specified group differences
         • Advantage over approach retaining only these groups is likely limited.
      ii. Can do more exotic things like confidence interval for extreme values of means.

4. Parametric Bootstrap:
   a. Replace sampling from sample by sampling from a parametric distribution whose mean and variance match the sample.

ST: 1.3

K. Related technique: Jackknife

1. Suppose that bias of estimator \( T \) of \( \theta \) based on \( n \) observations is \( a/n + b/n^2 + O(1/n^2) \).
   a. Ex., \( X_1, \ldots, X_n \) independent and identically distributed \( N(\mu, \sigma^2) \).
      i. MLE for \( \sigma^2 \) is \( \hat{\sigma}^2 = \sum_{j=1}^n (X_j - \bar{X})^2 / n \).
      ii. \( n\hat{\sigma}^2/(n-1) \) is unbiased estimator of \( \sigma^2 \).
      iii. \( E_{\sigma^2}[\hat{\sigma}^2] = (n-1)\sigma^2/n = \sigma^2 - \sigma^2/n \).
   b. Ex., \( W \) is an unbiased estimate of \( \omega \), \( T = g(W) \), \( \theta = g(\omega) \), \( Var[W] < \infty \).
   c. \( g(W) \approx g(\omega) + g'(\omega)(W-\omega) + g''(\omega)(W-\omega)^2/2 + g'''(\omega)(W-\omega)^3/6 + g''''(\omega)(W-\omega)^4/24 \).
   d. \( E[g(W)] \approx g(\omega) + g'(\omega)Var[W] / \| f \|^2 + g''(\omega)E[(W-\omega)^2] / 6 + g'''(\omega)E[(W-\omega)^3] / 24 + g''''(\omega)E[(W-\omega)^4] / 24 \).
5. Example where jackknife fails
   a. Assume that \( T_n \) acts sort of like mean
   b. If \( T_n \) really were a mean, then
      \[ \bar{T}_{n,i} = nT_n - (n-1)\bar{T}_{n-1,i} \]
      would be the observations that make it up.
   c. Then \( \text{Var}[\bar{T}_{n,i}] = n\text{Var}[T_n] \), and \( \bar{T}_{n,i} \) iid.
   d. Estimate \( \text{Var}[T_n] \) by \( \frac{n-1}{n} \sum_{i=1}^{n}(\bar{T}_{n,i} - \frac{1}{n} \sum_{j=1}^{n}\bar{T}_{n,j})^2 \).

4. Under some more restrictive conditions, can also use this to estimate variance.
   a. Expansion of expectation fails for rank-based estimators
   b. Ex., median for continuous data
      \[ \hat{T}_n = \frac{n+1}{4n}X_{((n-1)/2)} + \frac{n+1}{4n}X_{((n+1)/2)} + \frac{n-1}{2n}X_{((n+1)/2)} \]
   c. Then bias of \( \hat{T}_n - B \) is \( O(1/n^2) \).

3. Let \( T_n = \sum_{i=1}^{n}\bar{T}_{n-1,i}/n \).
   a. Then bias of \( \bar{T}_n \) is \( \approx a/(n-1) \).
   b. Let \( B = (n-1)(\bar{T}_n - T_n) \) be an estimate of the bias.
      i. Why? \( \text{E}[B] \approx (n-1)(\theta + a/(n-1) - \theta - a/n) = a(1 - a(n-1)/n) = a/n \).
      ii. Then bias of \( T_n - B \) is \( O(1/n^2) \).

2. Let \( \bar{T}_{n-1,i} \) be the estimator based on the sample of size \( n-1 \) with observation \( i \) omitted.

   d. \( n \) odd:
      i. \( T_n = X_{((n+1)/2)} \).
      ii. \( \bar{T}_{n-1,i} = \begin{cases} 
                     (X_{((n-1)/2)} + X_{((n+1)/2)})/2 & \text{if } i > (n+1)/2 \\
                     (X_{((n+1)/2)} + X_{((n+1)/2)})/2 & \text{if } i < (n+1)/2, \\
                     (X_{((n+1)/2)} + X_{((n-1)/2)})/2 & \text{if } i = (n+1)/2 
                    \end{cases} 
                       \text{with removed observation from ordered observations.} 
   v. Bias estimate is \( B = (n-1)\bar{T}_{n-1} - T_n = (n+1)(1/4X_{((n-1)/2)} + 1/4X_{((n+3)/2)} - 1/2X_{((n+1)/2)})/n \).

6. Works better for smooth functions of data like \( \alpha \) trimmed mean
   a. that is, mean of observation after smallest proportion \( \alpha \) removed, and largest proportion \( \alpha \) removed.