Bootstrap (continued)

- R live coding and debugging
- Parametric bootstrap
- Residual resampling, basic exchangeability
- Randomisation/permutation tests

Parametric bootstrap

- If the data size is small, basic Boostrap that resamples with replacement from the raw data has very little information.
- An alternative is to sample entirely new data from the model that was estimated on the whole data set.

Parametric Bootstrap sampling

Let $Y = \{(y_i, x_i), i = 1, \ldots, N\}$ be a data collection with response values y_i and predictors/covariates x_i, such that the conditional data density function is $p(y_i|x_i, \theta)$. Let $\widehat{\theta}$ be the maximum likelihood estimate of θ.

- Define the *Bootstrap sample* $Y^{(j)}$ by drawing N pairs $(y_i^{(j)}, x_i)$, where for each x_i, $y_i^{(j)}$ is drawn from $p(y_i|x_i, \widehat{\theta})$.
- Repeat this procedure for $j = 1, \ldots, J$, with $J \gg 1$.

Depending on the problem, the x_i values can either be kept fixed to their values in the original data set, resampled with replacement, or simulated from some generative model.

Residual resampling

- We can relax model assumptions, e.g. if we don't trust a Gaussian assumption for regression residuals.
- Instead of resampling the raw data, we resample the model *residuals*.

Residual resampling in regression models

Let $Y = \{(y_i, x_i), i = 1, \ldots, N\}$ be a data collection with response values y_i and predictors/covariates x_i, such that $\mu_i = \mathsf{E}(y_i | x_i, \theta)$ is the conditional expectation in a regression model. Let $\widehat{\theta}$ be the maximum likelihood estimate of θ, and $\widehat{\mu}_i = \mathsf{E}(y_i | x_i, \widehat{\theta})$.

- Define the residuals $r_i = y_i - \widehat{\mu}_i$, and construct a *residual Bootstrap sample* $Y^{(j)}$ by drawing N pairs $(y_i^{(j)}, x_i)$, where for each x_i, $y_i^{(j)} = \widehat{\mu}_i + r_i^{(j)}$, where $r_i^{(j)}$ is drawn with replacement from the empirical distributon of $\{r_1, r_2, \ldots, r_N\}$.
- Repeat this procedure for $j = 1, \ldots, J$, with $J \gg 1$.

For x_i, the same options as for fully parametric Bootstrap are available.

- This is a simple example of an *exchangeability* assumption; we no longer assume Gaussianity, but we do assume that all the residuals come from somecommon, but unknown, distribution; the individual residuals are (probabilistically) indistinguishable.

Randomisation/permutation tests

When assessing differences between two statistical populations, the hypotheses often take the form of some kind of *exchangeability structure*.

Exchangeability test example

- Let $Y_A = \{y_1^A, \dots, y_{N_A}^A\}$ and $Y_B = \{y_1^B, \dots, y_{N_B}^B\}$, and we want to test the hypotheses

$$H_0 : \text{The } A \text{ and } B \text{ come from the same distribution}$$
$$H_1 : \text{The } A \text{ and } B \text{ do not come from the same distribution}$$

- Given a test statistic $T(Y_A, Y_B)$, such as $\overline{y_\cdot^A} - \overline{y_\cdot^B}$, we need the distribution of T under H_0.
- Under H_0, the joint sample $Y_{A \cup B} = \{y_1^A, \dots, y_{N_A}^A, y_1^B, \dots, y_{N_B}^B\}$ is a collection of exchangeable variables.
- Each random permutation of $Y_{A \cup B}$ has the same distribution as $Y_{A \cup B}$, under H_0 (but not under H_1).

Randomisation/permutation tests

Assume that large test statistics T indicate deviations from H_0.

Permutation tests

- For $j = 1, \ldots, J$, draw $Y_{A\cup B}^{(j)}$ as a random permutation of $Y_{A\cup B}$, and split the result into subsets $Y_A^{(j)}$ and $Y_B^{(j)}$ of size N_A and N_B, respectively.
- Compute the test statistics $T^{(j)} = T(Y_A^{(j)}, Y_B^{(j)})$.
- The average $\frac{1}{J}\sum_{j=1}^{J} \mathbb{I}\{T^{(j)} \geq T(Y_A, Y_B)\}$ is an unbiased estimator of the p-value w.r.t. T for the hypothesis $H_0:$ the elements of Y_A and Y_B are mutually exchangeable.
- If $N_A + N_B$ is small, the limit $J \to \infty$ can be obtained by using all possible permutations instead of independent random permutations.

- In other exchangeability situations, similar tests can be designed.

Statistical Computing summary

- R data structures, indexing
- Structured programming
- Numerical maximum likelihood estimation
- Proper scoring functions for assessing predictions
- Floating point computation error analysis
- Cross validation
- Bootstrap
- Permutation tests