Bootstrapping

- different types: parametric (today), nonparametric (not covered)

  semiparametric (not covered)

- Monte Carlo requires us to know the underlying distribution under H_0
  - not always possible if H_0 is composite (i.e. many possibilities exist)

→ Utilize Parametric Bootstrap to estimate underlying distribution

Ex: Test for independence

You roll 2 4-sided dice and want to see if there is a correlation.

Data:

1st \ 2nd	1	2	3	4	Total
1	15	28	30	12	85
2	22	41	10	8	81
3	30	42	28	19	119
4	31	20	10	10	71
Total	98	131	78	49	356

χ^2-test for independence

$$\chi^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

i = row, j = column

$\chi^2 \approx 29.802$

$df = (4-1)(4-1) = 9$

p-value ≈ 0.0005

Estimate distribution of each roll by taking row totals and dividing by sample size

$$\hat{p}_1 = \left[\frac{85}{356}, \frac{81}{356}, \frac{119}{356}, \frac{71}{356}\right]$$

← estimated probability of the first roll resulting in a 4

$$\hat{p}_2 = \left[\frac{98}{356}, \frac{131}{356}, \frac{78}{356}, \frac{49}{356}\right]$$