STATISTISCHE METHODEN DER DATENANALYSE

Excersise Sheets

Physik

UNI WIEN

23. November 2023

1 Sheet

1.1 Bernoulli

1.1.1 Clopper and Pearson confidence interval

A Bernoulli experiment is repeated $n = 200$ times with $k = 121$ sucesses. Calculate the symmetric 95% interval for the parameter p. The Interval boundaries can be calculated with the inverse beta distribution.

$$G_1(k) = \beta\left(\frac{\alpha}{2}; k, n-k+1\right) = 0.534 \tag{1.1}$$

$$G_2(k) = \beta\left(\frac{1-\alpha}{2}; k+1, n-k\right) = 0.673 \tag{1.2}$$

1.1.2 Approximation by normal distribution (bootstrap and robust)

Estimate p:

$$\hat{p} = \frac{k}{n} = \frac{121}{200} \tag{1.3}$$

With that estimate σ

$$\sigma[\hat{p}] = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \tag{1.4}$$

$$= \sqrt{\frac{9559}{8 \cdot 10^6}} \approx 0.035 \tag{1.5}$$

$$z_{1-\frac{\alpha}{2}} = f_{\text{norm}}\left(1 - \frac{\alpha}{2}\right) = 0.248 \tag{1.6}$$

Now the interval boundaries are for the bootstrap method:

$$G_1(k) = \hat{p} - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \approx 0.591 \tag{1.7}$$

$$G_1(k) = \hat{p} + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \approx 0.619 \tag{1.8}$$

and for the robust method

$$G_1(k) = \hat{p} - z_{1-\frac{\alpha}{2}} \frac{1}{2\sqrt{n}} \approx 0.585 \tag{1.9}$$

$$G_1(k) = \hat{p} + z_{1-\frac{\alpha}{2}} \frac{1}{2\sqrt{n}} \approx 0.625 \tag{1.10}$$

1.1.3 Agresti-Coull

$$G_1(k) \approx 0.585 G_2(k) \approx 0.624 \tag{1.11}$$

1.2