00137960: Statistical Thinking
Homework 2

1. (Programming) Suppose that $x_i \sim N(\mu, 1)$ independent, $Z_i = \sum_{j=1}^{i} x_j/\sqrt{i}$, and we wish to test $H_0 : \mu = 0$ vs. $H_1 : \mu > 0$. Show that the stopping rule "either Z_{20} or Z_{30} exceeds 1.645" has probability 0.074 of rejecting H_0 if H_0 is true.

2. For an i.i.d. sample of size n from the bivariate normal distribution with correlation ρ, use the delta method to show that the sample correlation coefficient $\widehat{\rho}$ has the asymptotic distribution

$$\sqrt{n}(\widehat{\rho} - \rho) \to_d N\big(0, (1-\rho^2)^2\big).$$

3. In the ulcer surgery example in Table 4.1 of Efron and Hastie, construct a test for the null hypothesis that the two treatments are equally effective, conditional on the marginals. Evaluate your test statistic on the data and conclude.

4. Find the Fisher information for the Cauchy distribution with density

$$f_\theta(x) = \frac{1}{\pi} \frac{1}{1 + (x-\theta)^2}.$$

5. If $\boldsymbol{\mu} \sim N_p(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ and $\boldsymbol{x} \mid \boldsymbol{\mu} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, show that the posterior distribution of $\boldsymbol{\mu}$ is

$$\boldsymbol{\mu} \mid \boldsymbol{x} \sim N_p\big(\boldsymbol{\mu}_0 + (\boldsymbol{\Sigma}_0^{-1} + \boldsymbol{\Sigma}^{-1})^{-1}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_0), (\boldsymbol{\Sigma}_0^{-1} + \boldsymbol{\Sigma}^{-1})^{-1}\big).$$

6. For a one-parameter exponential family $f_\alpha(x) = e^{\alpha y - \psi(\alpha)} f_0(x)$, verify by direct differentiation the following relationships between derivatives of $\psi(\alpha)$ and the mean μ, variance σ^2, and third and fourth central moments μ_3, μ_4 of y:
   (a) $\psi'(\alpha) = \mu$;
   (b) $\psi''(\alpha) = \sigma^2$;
   (c) $\psi^{(3)}(\alpha) = \mu_3$;
   (d) $\psi^{(4)}(\alpha) = \mu_4 - 3(\sigma^2)^2$.

7. (Programming) Fit the seven-parameter exponential family with $f_0(x) \equiv 1$ and $y = (x, x^2, \ldots, x^7)$ to the glomerular filtration rate data (available at https://hastie.su.domains/CASI_files/DATA/gfr.txt) and reproduce Figure 5.7 of Efron and Hastie.