J Kohonen

6B Hypothesis testing

This is the last set of exercises. There are no more home problems.

Class problems

6B1 (Nail factory) A factory is producing nails, with the aim that the average length should be 10.00 cm. Because of imprecisions in the manufacturing process, each produced nail has a slightly random length, according to a normal distribution with an unknown mean μ and an unknown standard deviation σ. For quality control, a sample of 120 nails were inspected. In the sample, the average was 10.08 cm and the sample standard deviation was 0.40 cm.

Because of the large sample size, we work as if the standard deviation of the nail-generating process is $\sigma = 0.40$ (i.e. equals the sample standard deviation). This is Ross's "known variance" situation (§7.3 and 8.3.1).

(a) Compute a 95% confidence interval for the unknown mean.

(b) Compute a 99% confidence interval for the unknown mean.

(c) Calculate the p-value for a hypothesis test with the null hypothesis $\mu = 10.00$, against the alternative hypothesis that $\mu \neq 10.00$.

(d) Using the p-value from (b), report whether the null hypothesis is accepted or rejected at significance level $\alpha = 0.05$.

(e) Using the p-value from (b), report whether the null hypothesis is accepted or rejected at significance level $\alpha = 0.01$.

(f) (Optional) Redo (a)–(e), but now take into account the uncertainty about σ (see Ross's "unknown variance" situation §7.3.1 and 8.3.2). That is, instead of using the standard normal distribution, use Student's t distribution with parameter $n - 1 = 119$. Compare the numerical results to the previous ones.

The CDF and its inverse for the t distribution can be computed in R with `pt` and `qt`, and in Matlab with `tcdf` and `tinv`.

6B2 (Peas) Gregor Mendel (1822–1884) laid the foundation for modern genetics by studying how certain features of pea plants are inherited. Some of his experiments led to the theory that when pea plants are cross-bred in a particular way, then 75% of the offspring plants produce yellow peas, and 25% produce green peas.

In order to test this theory, a geneticist cross-breeds 80 plants using the same procedure, and observes that 56 of them produce yellow peas. Is this observation compatible with Mendel's theory? Formulate and perform a (two-sided) hypothesis test. Let the null hypothesis be that each plant has probability $\theta = 0.75$ of producing yellow peas (independently of the other plants). The alternative hypothesis is that $\theta \neq 0.75$. Apply the significance level 5%.

You can use either the exact binomial test (which exists in two versions; see Lecture 6A and Ross 8.6), or the normal approximation.

Note that in a binary model, one might commonly use "p" to denote the probability parameter (for a plant to produce yellow peas), but here we use θ to avoid confusion with the p-value, which is a probability associated with the hypothesis test.