18.650. Statistics for Applications
Fall 2016. Problem Set 5

Due Friday, Oct. 14 at 12 noon

Problem 1 Hypotheses testing and confidence intervals

Consider a sample $X_1, \ldots, X_n \overset{i.i.d.}{\sim} \mathcal{P}oiss(\lambda)$, for some unknown $\lambda > 0$.

1. Based on the sample, propose a confidence interval for λ with asymptotic level $1-\alpha$, for some fixed $\alpha \in (0,1)$ (prove your answer). Denote by I your confidence interval.
2. Consider the following hypotheses, where λ_0 is a given positive number:

$$H_0 : \text{"}\lambda = \lambda_0\text{"} \text{ vs. } H_1 : \text{"}\lambda \neq \lambda_0\text{"}.$$

Using the confidence interval I, propose a test with asymptotic level α for these hypotheses.

Problem 2 Comparing two means

Consider two measuring instruments that are used to measure the intensity of some electromagnetic waves. An engineer wants to check if both instruments are calibrated identically, i.e., if they will produce identical measurements for identical waves. To do so, the engineer does n_1 independent measurements of the intensity of a given wave using the first instrument, and n_2 measurements on the same wave using the second instrument. The integers n_1 and n_2 may not be equal because, for instance, one instrument may be more costly than the other one, or may produce measurements more slowly. The measurements are denoted by $X_1, \ldots, X_{n_1}$ for the first instrument and by $Y_1, \ldots, Y_{n_2}$ for the second one. Intrinsic defects of the instruments will lead to measurement errors, and it is reasonable to assume that $X_1, \ldots, X_{n_1}$ are iid Gaussian and so are $Y_1, \ldots, Y_{n_2}$. If the two instruments are identically calibrated, the X_i's and the Y_i's should have the same expectation but may not have the same variance, since the two instruments may not have the same precision.

Hence, we assume that $X_1, \ldots, X_{n_1} \overset{i.i.d.}{\sim} \mathcal{N}(\mu_1, \sigma_1^2)$ and $Y_1, \ldots, Y_{n_2} \overset{i.i.d.}{\sim} \mathcal{N}(\mu_2, \sigma_2^2)$, where $\mu_1, \mu_2 \in \mathbb{R}$ and $\sigma_1^2, \sigma_2^2 > 0$, and that the two samples are independent of each other. We want to test whether $\mu_1 = \mu_2$.

1. Recall the expression of the maximum likelihood estimators for (μ_1, σ_1^2) and for (μ_2, σ_2^2). Denote these estimators by $(\hat{\mu}_1, \hat{\sigma}_1^2)$ and $(\hat{\mu}_2, \hat{\sigma}_2^2)$.

2. Recall the distribution of $\frac{n_1\hat{\sigma}_1^2}{\sigma_1^2}$ and of $\frac{n_2\hat{\sigma}_2^2}{\sigma_2^2}$.

3. What is the distribution of $\frac{n_1\hat{\sigma}_1^2}{\sigma_1^2} + \frac{n_2\hat{\sigma}_2^2}{\sigma_2^2}$?

4. Let $\Delta = \hat{\mu}_1 - \hat{\mu}_2$. What is the distribution of Δ ?

5. Consider the following hypotheses:

$$H_0 : \text{"}\mu_1 = \mu_2\text{"} \text{ vs. } H_0 : \text{"}\mu_1 = \mu_2\text{"}.$$

   Here and in the next question we assume that $\sigma_1^2 = \sigma_2^2$. Based on the previous questions, propose a test with non asymptotic level $\alpha \in (0, 1)$ for H_0 against H_1.

6. Assume that 10 measurements have been done for both machines. The first instrument measured 8.43 in average with sample variance 0.22 and the second instrument measured 8.07 with sample variance 0.17. Can you conclude that the calibrations of the two machines are significantly identical at level 5% ? What is, approximately, the p-value of your test ?

Problem 3 Implicit hypotheses testing

Based on a sample of i.i.d. Gaussian random variables $X_1, \ldots, X_n$ with mean μ and variance σ^2, propose a test with asymptotic level 5% for the hypotheses

$$H_0 : \text{"}\mu > \sigma\text{"} \text{ vs. } H_1 : \text{"}\mu \leq \sigma\text{"}.$$

What is the p-value of your test if the sample has size $n = 100$, the sample average is 2.41 and the sample variance is 5.20 ? If the sample size is $n = 100$, the sample average is 3.28 and the sample variance is 15.95 ? In the latter case, do you reject H_0 at level 5% ? At level 10% ?