EXAM Probability theory and statistical inference I , 2ST065 (7.5 hp).
Wednesday 19/10 2016, 8.00 - 13.00.
Examiner: Patrik Andersson.

Allowed tools.
– Formulae for the course Probability Theory and Statistical Inference
– Math Handout (by Lars Forsberg)
– Pocket calculator.
– Physical dictionary (or word-list).

Notes in the permitted aids are not allowed. If you feel that something in the wording of the problem is unclear, write under what assumptions you are solving the problem. After turning in your test, you may keep the test-pages with the question-statements. For the grade Pass, a score of at least 50% is required on the exam.

1. In a bucket, there are 15 balls of different color, 5 blue, 5 red and 5 green. The balls are, in addition to a having a color, numbered from 1 to 15. We draw 5 balls from the bucket, without replacement.
   (a) Find the probability of obtaining 2 or less blue balls. (10p)
   (b) Find the expected value of the sum of the numbers on the drawn balls. (10p)
2. A random variable, X, has a half-normal distribution if the density function is
$$f_X(x) = \frac{\sqrt{2}}{\sqrt{\sigma^2 \pi}} e^{-\frac{x^2}{2\sigma^2}},\ x > 0.$$
   It can be shown that $E[X] = \sqrt{\frac{2\sigma^2}{\pi}}$. The following observations were made on X:

Observation #	1	2	3	4	5
X	1.43	0.32	0.75	1.37	1.71

   (a) Given the above observations, find the maximum likelihood estimate of σ^2. (15p)
   (b) Find the maximum likelihood estimate of $E[X]$. (5p)
3. Consider the following joint density of X and Y,
$$f_{X,Y}(x, y) = Ce^{-(x+y)},\ x > 0,\ y > 0.$$
   (a) Find the constant C, such that this is a true density. (5p)

(b) Find $E[X]$. (5p)

(c) Find $\mathrm{Var}(Y)$. (5p)

(d) Find $\mathrm{Cov}(X, Y)$. (5p)

4. Two samples, from two populations, were made. The following table summarize the samples

	Sample mean	Sample std. dev.	Number of observations
Sample 1	1.70	1.22	10
Sample 2	0.65	1.72	10

(a) Test on the 5% level if the variance in the two populations are the same. Also state what assumptions are needed for the test. (10p)

(b) Test on the 5% level if the population mean in the two populations are the same. Also state what assumptions are needed for the test. (10p)

Solutions Probability theory and statistical inference I , 2ST065 (7.5 hp).
Wednesday 19/10 2016
Examiner: Patrik Andersson.

1. (a) Let X denote the number of blue balls. Then $X \sim \mathsf{Hyp}(15, 5, 5)$. So,

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$
$$= \frac{\binom{5}{0}\binom{10}{5}}{\binom{15}{5}} + \frac{\binom{5}{1}\binom{10}{4}}{\binom{15}{5}} + \frac{\binom{5}{2}\binom{10}{3}}{\binom{15}{5}} = \ldots = \frac{2502}{3003} \approx 0.82.$$

(b) Now, let Y_1 to Y_5 denote the number on balls 1 to 5. Note that these random variables have the same distribution and that $E[Y_i] = 8$. Thus

$$E[Y_1 + \cdots + Y_5] = E[Y_1] + \cdots + E[Y_5] = 5 \cdot 8 = 40.$$

2. (a) Denoting the 5 observations by x_1 to x_5, the likelihood function is

$$L(\sigma^2) = \left(\frac{2}{\pi\sigma^2}\right)^{5/2} \prod_{i=1}^{5} e^{-\frac{x_i^2}{2\sigma^2}},$$

so that the log-likelihood function is

$$l(\sigma^2) = C - \frac{5}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{5} x_i^2.$$

We differentiate to find the maximum,

$$\frac{d}{d\sigma^2} l(\sigma^2) = -\frac{5}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{5} x_i^2 = 0.$$

With the solution,

$$\hat{\sigma}^2 = \frac{1}{5}\sum_{i=1}^{5} x_i^2 \approx 1.50,$$

which is the maximum likelihood estimate.

(b) By the invariance property of the MLE we get the MLE of $E[X]$ to be,

$$\widehat{E[X]} = \sqrt{\frac{2\hat{\sigma}^2}{\pi}} \approx \sqrt{\frac{2 \cdot 1.50}{\pi}} \approx 0.98.$$

3. Since the joint density can be factorized we see immediately that X and Y are independent and they are both distributed according to $\mathsf{Exp}(1)$.
   (a) By looking in the table we see that $C = 1$.
   (b) Again, the table gives $E[X] = 1$.
   (c) The table gives $\text{Var}(X) = 1$.
   (d) By independence, $\text{Cov}(X, Y) = 0$.
4. (a) Denoting the variances in the two populations σ_1^2 and σ_2^2, we shall test $\sigma_1^2 = \sigma_2^2$ against $\sigma_1^2 \neq \sigma_2^2$. The test statistic is

$$F = \frac{s_2^2}{s_1^2} = \left(\frac{1.72}{1.22}\right)^2 \approx 1.98,$$

   which, if the null-hypothesis is true, is an observation from the $F(9, 9)$ distribution. We thus compare this to $F_{0.025}(9, 9) = 4.03$. Since $1.98 < 4.03$ we accept $\sigma_1^2 = \sigma_2^2$. For this test to be valid we need to assume that both populations are normally distributed.

   (b) Now, denoting the means of the two populations by μ_1 and μ_2, we shall test $\mu_1 = \mu_2$ against $\mu_1 \neq \mu_2$. For this test we again need to assume normally distributed populations and also that $\sigma_1^2 = \sigma_2^2$, which is supported by (a).

   The pooled sample variance is

$$s_p^2 = \frac{9 \cdot (1.22)^2 + 9 \cdot (1.72)^2}{9 + 9} \approx 2.22.$$

   The test statistic, which under the null hypothesis is t(18) distributed, is

$$t = \frac{1.70 - 0.65}{\sqrt{2.22/5}} \approx 1.58.$$

   The critical value is $t_{0.025}(18) = 2.101$. Since $1.58 < 2.101$ we accept that $\mu_1 = \mu_2$.