EXAM Probability theory and statistical inference I , 2ST065 (7.5 hp).
Wednesday 16/3 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, state 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 75% is required on the Pass part of the exam. For the grade *Pass with distinction*, a score of at least 75% is required on the Pass part of the exam *and* a score of at least 50% on the Pass with distinction part of the exam.

Pass part

1. (20p) A motorcycle has two brakes, front and back. In order to pass the yearly inspection, both brakes have to work properly.

   This particular motorcycle is in bad condition so that on a given day, the probability that the front brake will work is 0.9 and the probability that the front brake works but the back brake malfunctions is 0.2.

   What is the probability of passing the inspection?

2. Consider the density
$$f_X(x) = \frac{3x^2}{a^3}, 0 < x < a, a > 0.$$
   The following observations on X are made

Obs.	1	2	3	4	5
X	1.7908	7.4022	1.0706	1.0386	1.6070

   (a) Find the method of moments estimate of a. (10p)

   (b) Is the estimator unbiased? Motivate! (5p)

   (c) Is the estimator consistent? Motivate! (5p)

3. A sample of size 10 was taken from a $\mathsf{N}(\mu, \sigma^2)$ distribution. The sample mean and standard deviation was $\bar{x} = 0.33$ and $s = 0.96$.

   (a) Test on the 5%-level if $\mu = 0$ against $\mu \neq 0$. (5p)

   (b) What is the p-value of this test? (5p)

   (c) Calculate the Type II-error probability, assuming that $\mu = 0.1$. (10p)

4. Consider the the bivariate density $f_{X,Y}(x, y) = Cx(x + y),\ 0 < x < 1,\ 0 < y < 1$.

   (a) Find C such that this a true density. (7p)

   (b) Calculate $P(X \leq 0.5|Y > 0.5)$. (6p)

   (c) Calculate $P(X \leq 0.5|Y = 0.5)$. (6p)

Pass with distinction part

5. (20p) Let $X \sim \mathsf{N}(\mu, \sigma^2)$. Find the density of $Y = e^X$.

6. (20p) An urn contains 11 balls. The balls are either red or blue. Five balls were sampled from the urn without replacement; 2 red balls and 3 blue balls were obtained.

   Now, all the balls are put back in the urn and then 5 more balls are drawn, this time with replacement. Find the maximum likelihood estimate of the probability that at least 1 of the the drawn balls is blue.

Solutions Probability theory and statistical inference I , 2ST065 (7.5 hp).
Wednesday 16/3 2016
Examiner: Patrik Andersson.

1. Let the notation be $F = \{\text{front brake works}\}$ and $B = \{\text{back brake works}\}$. We are then looking for

$$P(F \cap B) = P(F) - P(F \cap \bar{B}) = 0.9 - 0.2 = 0.7.$$

2. (a) We first calculate the expected value of the distribution.

$$\mu = \int_0^a \frac{3x^3}{a^3} dx = \frac{3}{a^3} \left[\frac{x^4}{4} \right]_0^a = \frac{3a}{4}.$$

Thus, $a = 4\mu/3$. We estimate μ by the usual moment estimator $\hat{\mu} = 1/n \sum_i x_i$ and get the estimtor,

$$\hat{a} = \frac{4}{3} \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Using the observations we get the estimate $\hat{a} = 3.44$.

(b) The expected value of the estimator is

$$\mathsf{E}\,[\hat{a}] = \mathsf{E}\left[\frac{4}{3} \frac{1}{n} \sum_{i=1}^{n} X_i \right] = \frac{4}{3} \frac{1}{n} \sum_{i=1}^{n} \mathsf{E}\,[X_i] = \frac{4}{3} \frac{1}{n} \sum_{i=1}^{n} \frac{3a}{4} = a.$$

Since the expected value of the estimator is equal to the parameter, the estimator is unbiased.

(c) We know that an unbiased estimator is consistent if the variance decreases to 0 as the sample size increases. It is therefore enough to note that

$$\mathrm{Var}\,(\hat{a}) = \mathrm{Var}\left(\frac{4}{3} \frac{1}{n} \sum_{i=1}^{n} X_i \right) = \frac{16}{9} \frac{1}{n} \mathrm{Var}(X_i) \to 0, \text{ as } n \to \infty.$$

Where we used that $\mathrm{Var}(X_i)$ is a finite number, the exact value of which is not important for the argument.

3. The test statistics is

$$t = \frac{0.33}{0.96/\sqrt{10}} = 1.08.$$

Under the assumption that $\mu = 0$ this should be $\mathsf{t}(9)$ distributed.

(a) In order to do a double-sided test on the 5%-level we should compare this to $t_{0.025}(9) = 2.262$. We see that $t = 1.08 < 2.262$ and we therefore can not reject that $\mu = 0$.

(b) Since $t_{0.1}(9) = 1.382$ we can conclude that the p-value is larger than 0.2.

(c) A Type II-error is to not reject $\mu = 0$ even though $\mu \neq 0$; in our case $\mu = 0.1$. The Type II-error probability for the test on the 5%-level is therefore

$$P_{\mu=0.1}\left(-2.262 \leq \frac{\bar{X}}{s/\sqrt{10}} \leq 2.262\right)$$

This probability is difficult to calculate, so let us continue under the assumption that σ is known to be 0.96. Then

$$\begin{aligned}
&P_{\mu=0.1}\left(-2.262 \leq \frac{\bar{X}}{1.96/\sqrt{10}} \leq 2.262\right)\\
=&P_{\mu=0.1}\left(-2.262 - \frac{0.1}{1.96/\sqrt{10}} \leq \frac{\bar{X}-0.1}{1.96/\sqrt{10}} \leq 2.262 - \frac{0.1}{1.96/\sqrt{10}}\right)\\
=&P\left(-2.59 \leq Z \leq 1.93\right) = P(Z \leq 1.93) + P(Z \leq 2.59) - 1\\
=&0.9732 + 0.9952 - 1 = 0.97
\end{aligned}$$

4. (a) Since densities should integrate to 1,

$$\frac{1}{C} = \int_0^1 \int_0^1 x(x+y)dxdy = \int_0^1 \left[\frac{x^3}{3} + \frac{x^2}{2}y\right]_0^1 dy \tag{1}$$

$$= \int_0^1 (\frac{1}{3} + \frac{y}{2})dy = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}. \tag{2}$$

Thus, $C = 12/7$.

(b)

$$P(X \leq 0.5|Y > 0.5) = \frac{P(X \leq 0.5, Y > 0.5)}{P(Y > 0.5)}.$$

For the numerator we calculate

$$\int_{0.5}^1 \int_0^{0.5} \frac{12}{7}x(x+y)dxdy = \ldots = \frac{5}{48},$$

for the denominator

$$\int_{0.5}^1 \int_0^1 \frac{12}{7}x(x+y)dxdy = \ldots = \frac{17}{48},$$

so that $P(X \leq 0.5|Y > 0.5) = 5/17$.

(c) We start by finding

$$f_{X|Y}(x, 0.5) = \frac{f_{X,Y}(x, 0.5)}{f_Y(0.5)} = \frac{12}{7}x(x + 0.5),$$

where we used that

$$f_Y(y) = \int_0^1 \frac{12}{7}x(x + y)dx = \frac{12}{7}\left[\frac{x^3}{3} + \frac{x^2}{2}y\right]_0^1 = \frac{12}{7}\left(\frac{1}{3} + \frac{y}{2}\right),$$

so that $f_y(0.5) = 1$. We get that

$$P(X \leq 0.5|Y = 0.5) = \int_0^{0.5} f_{X|Y}(x, 0.5)dx = \frac{12}{7}\int_0^{0.5} x(x + 0.5)dx$$
$$= \frac{12}{7}\left[\frac{x^3}{3} + \frac{x^2}{4}\right]_0^{0.5} = \frac{12}{7}\left(\frac{1}{3 \cdot 8} + \frac{1}{4 \cdot 4}\right) = \frac{5}{28}.$$

5. X has the density

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \ x \in \mathbb{R}.$$

For Y we have that,

$$F_Y(y) = P(Y \leq y) = P(e^X \leq y) = P(X \leq \log y) = F_X(\log y).$$

Further,

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X(\log y) = \frac{dF_X(\log y)}{d(\log y}\frac{d\log y}{dy}$$
$$= f_X(\log y)\frac{1}{y} = \frac{1}{y\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(\log y-\mu)^2}, \ x \in \mathbb{R}.$$

6. Let X denote the number of blue balls obtained in the first 5 ball sample. Then $X \sim \mathsf{Hyp}(11, 5, r)$, were r is the unknown number of blue balls. The MLE of r is that which gives the largest $P(X = 3)$. We thus calculate this probability for the different possibilites of r, noting, because $X = 3$, that $3 \leq r \leq 9$

r	$P(X = 3)$
3	0.06
4	0.18
5	0.32
6	0.43
7	0.45
8	0.36
9	0.18

This means that the MLE of r is $\hat{r} = 7$.

Now we calculate, assuming we know r, the probability of obtaining at least 1 blue ball in the second draw. Let Y be the number of blue balls drawn, then $Y \sim \mathsf{Bin}(5, r/11)$. Thus,

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - \left(1 - \frac{r}{11}\right)^5 = g(r),$$

which is a function $g(r)$. The maximum likelihood estimate of $g(r)$ is $\widehat{g(r)} = g(\hat{r})$. That is,

$$\widehat{g(r)} = g(7) = 1 - \left(1 - \frac{7}{11}\right)^5 = 0.9936.$$