EXAM Probability theory and statistical inference I , 2ST065 (7.5 hp).
Friday 19/2 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. In three bowls, the number of red and black balls are as follows

Bowl	# Red	# Black
1	2	2
2	3	1
3	4	0

First, one of bowl 1, 2 and 3 is selected with probabilities 0.3, 0.3 and 0.4 respectively. After that 2 balls are drawn from the bowl, without replacement.

(a) Find the probability that at least one black ball is drawn. (10p)

(b) Given that at least one black ball is drawn, what is the probability that it was bowl 1 that was selected? (10p)

2. (20p) The number of points on an exam for a randomly selected student can be assumed to be distributed as $\mathsf{N}(50, 100)$. The examiner would like to set the number of points required for a Pass such that on average 70% of the students pass the exam. How many points should the examiner require for a Pass?

3. Let $X \sim \mathsf{U}(-a, a)$, $a > 0$. Assume that you have a sample of size n denoted x_i, $i = 1 \ldots n$ from this distribution.

(a) Find the method of moments estimator of a. (7p)

(b) Consider the alternative estimator $\hat{a} = \frac{2}{n}\sum_{i=1}^{n}|x_i|$. Show that this estimator is unbiased. (7p)

(c) Is this estimator consistent? (6p)

Hint: Note that for $a > 0$, $\int_{-a}^{a}|x|dx = \int_{-a}^{0} -xdx + \int_{0}^{a} xdx$.

4. (20p) Assume that $X_i \sim \mathsf{N}(\mu_x, \sigma_x^2)$ and $Y_i \sim \mathsf{N}(\mu_y, \sigma_y^2)$. Samples of size 50 were taken from each distribution and the following was calculated, $\bar{x} = 0.4394$, $\bar{y} = -0.3775$, $s_x^2 = 1.160$ and $s_y^2 = 3.641$.

   Test on the 1%-level if the two variances are equal against the alternative that σ_y^2 is larger.

Pass with distinction part

5. Consider the following bivariate density,

$$f_{X,Y}(x,y) = cxy^2,\ 0 \le x \le 1 + y \le 2.$$

   (a) Find c such that $f_{X,Y}(x,y)$ is a density. (10p)

   (b) Calculate $P(0.5 \le X \le 1.5, Y \ge 0)$ (10p)

6. (20p) In a building all the 10 lights were replaced and new fluorescent lamps were installed. The life time of the lamps are assumed to be exponentially distributed and the manufacturer claims that the mean life time is 5 years.

   Already after one month the first lamp broke and we now wish to test whether the manufacturer has exaggerated the mean life time of the lamps.

   Perform the test and calculate the p-value.

   Hint: You may use that for $X_1, \ldots, X_n$ iid $\mathsf{Exp}(\beta)$, $\min\{X_1, \ldots, X_n\} \sim \mathsf{Exp}(\beta/n)$.

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

1. Let B_i denote the event that bowl number i was chosen and let X be the number of black balls drawn.

   (a) We will use the law of total probability and condition on which bowl was chosen and calculate the probability of the complement of *at least one ball drawn*, $P(X = 0)$

   $$\begin{aligned} P(X=0) &= P(X=0|B_1)P(B_1) + P(X=0|B_2)P(B_2) + P(X=0|B_3)P(B_3) \\ &= \frac{2}{4} \times \frac{1}{3} \times 0.3 + \frac{3}{4} \times \frac{2}{3} \times 0.3 + \frac{4}{4} \times \frac{3}{3} \times 0.4 = 0.6. \end{aligned}$$

   Thus $P(X \geq 1) = 1 - P(X = 0) = 0.4$.

   (b) Using Bayes' theorem

   $$P(B_1|X \geq 1) = \frac{P(X \geq 1)|B_1)P(B_1)}{P(X \geq 1)} = \frac{(1 - \frac{2}{4} \times \frac{1}{3}) \times 0.3}{0.4} = \frac{5}{8}.$$

2. Let $X \sim \mathsf{N}(50, 100)$ be the number of points on the exam for a random student. We want to find k such that $P(X > k) = 0.7$.

   $$P(X > k) = P\left(\frac{-50}{\sqrt{100}} > \frac{k-50}{\sqrt{100}}\right) = P\left(Z > \frac{k-50}{\sqrt{100}}\right) = 0.7,$$

   where $Z \sim \mathsf{N}(0, 1)$. Taking complement we write this as

   $$P\left(Z \leq \frac{k-50}{\sqrt{100}}\right) = 0.3,$$

   and therefore see that we should choose k such that $\frac{k-50}{\sqrt{100}} \approx -0.52$. We thus see that $k \approx 44.8$ and the limit can be set at 44 points.

3. (a) We have that $E[X] = 0$. Since this does not involve the parameter we want to estimate, we can not use it to find an estimator. We continue with the second moment, $E[X^2] = \mathrm{Var}(X) + E[X]^2 = a^2/3$. We will estimate this with $\hat{\mu}_2 = \frac{1}{n}\sum_{i=1}^n x_i^2$. The method of moment is then to estimate a by solving

   $$\hat{\mu}_2 = \frac{\hat{a}^2}{3},$$

   with solutions $\hat{a} = \pm\sqrt{3\hat{\mu}_2}$. Since we know that $a > 0$, the negative solution is not a good estimator. So the method of moment estimator is $\hat{a} = \sqrt{3\hat{\mu}_2} = \sqrt{\frac{3}{n}\sum_{i=1}^n x_i^2}$.

(b) To show that it is unbiased we should show that $E[\hat{a}] = a$.

$$\mathsf{E}\left[\hat{a}\right] = \mathsf{E}\left[\frac{2}{n}\sum_{i=1}^{n}|X_i|\right] = \frac{2}{n}\sum_{i=1}^{n}E\left[|X_i|\right].$$

So, it is enough to show that $E\left[|X_i|\right] = a/2$. This is true since,

$$\begin{aligned} E\left[|X_i|\right] &= \int_{-a}^{a}|x|\frac{1}{2a}dx = \frac{1}{2a}\left[\int_{-a}^{0}-xdx + \int_{0}^{a}xdx\right] \\ &= \frac{1}{2a}\left\{\left[\frac{-x^2}{2}\right]_{-a}^{0} + \left[\frac{x^2}{2}\right]_{0}^{a}\right\} = \frac{a^2}{2a} = \frac{a}{2}. \end{aligned}$$

(c) The estimator is unbiased, thus it is enough to show that the variance goes to 0 as the sample size increases.

$$\mathrm{Var}\left(\frac{2}{n}\sum_{i=1}^{n}|x_i|\right) = \{\text{independent sample}\} = \frac{4}{n}\mathrm{Var}\left(|X_i|\right) \to 0 \text{ as } n \to \infty,$$

since $\mathrm{Var}(|X_i|)$ is some finite number.

4. We wish to test $H_0:\ \sigma_x^2 = \sigma_y^2$ against $H_a:\ \sigma_x^2 < \sigma_y^2$. The test statistic is $F = s_y^2/s_x^2$ which is a sample from the $F(49, 49)$ distribution if H_0 is true. We will reject H_0 if $F > F_{0.01}(49, 49)$ on the 1% level.

   We get that $F = 3.641/1.160 \approx 3.14$ and that $F_{0.01}(49, 49) \approx F_{0.01}(40, 40) = 2.11$. We thus reject H_0 on the 1% level.

5. Let us calculate the integral

(a)

$$\begin{aligned} \int_{-1}^{1}\int_{0}^{1+y} xy^2dxdy &= \int_{-1}^{1} y^2\left[\frac{x^2}{2}\right]_0^{1+y}dy = \int_{-1}^{1}y^2\frac{(1+y)^2}{2}dy \\ &= \frac{1}{2}\int_{-1}^{1}y^2(y^2+2y+1)dy = \{\text{symmetry}\} = \int_{0}^{1}(y^4+y^2)dy = \frac{8}{15}. \end{aligned}$$

Thus we need $c = 15/18$ for $f_{X,Y}(x, y)$ to be a proper density.

(b)

$$\begin{aligned} P(0.5 \le X \le 1.5, Y \ge 0) &= \int_{0.5}^{1}\int_{0.5}^{1.5}\frac{15}{18}xy^2dxdy + \int_{0}^{0.5}\int_{0.5}^{1+y}\frac{15}{18}xy^2dxdy \\ &= \ldots = \frac{313}{512}. \end{aligned}$$

6. The life time of a lamp is $X_i \sim \mathsf{Exp}(\beta)$. We wish to test $H_0 : \beta = 5$ against $H_a : \beta < 5$. Our test statistic will be $Y =$ the time until the first lamp breaks. Under H_0,

$$Y = \min\{X_1, \ldots, X_{10}\} \sim \mathsf{Exp}(5/10) = \mathsf{Exp}(0.5).$$

We observed $Y = 1/12$ and will reject H_0 if $Y < k$, where k is such that $P_{H_0}(Y < k) = \alpha$. So the p-value is

$$\begin{aligned} \text{p-value} = P_{H_0}(Y < 1/12) &= \int_0^{1/12} \frac{1}{0.5} e^{-x/0.5} dx \\ &= 1 - e^{-(1/12)/0.5} \approx 0.15. \end{aligned}$$