Norwegian University of
Science and Technology

Department of Mathematical Sciences

Examination paper for **TMA4240 Statistics**

Academic contact during examination: Håkon Tjelmeland[a], Sara Martino[b]

Phone: [a]48 22 18 96 , [b]99 40 33 30

Examination date: November 30th 2017

Examination time (from–to): 09.00-13.00

Permitted examination support material: C

Tabeller og formler i statistikk, Akademika,
A specific basic calculator
Yellow stamped A5-sheet with personal hand written notes.

Other information:
All your answers should be justified and the hand-in material should contain calculations leading to your answer.

Language: English

Number of pages: 4

Number of pages enclosed: 0

Checked by:

Informasjon om trykking av eksamensoppgave	
Originalen er:	
1-sidig ☐	**2-sidig** ☒
sort/hvit ☒	**farger** ☐
skal ha flervalgskjema ☐	

Date Signature

Problem 1

Assume we throw two regular dices. For the two dices combined, let A denote the event that this results in at least six pips and let X denote the number of pips.

Find $P(A)$. Explain the interpretation of this number.

Find $\mathrm{E}[X]$. Explain the interpretation of this number.

Problem 2 The follow-up department

Two departments in a hospital in a larger city, which are both doing surgery, have a common follow-up department. We will call the two departments for department A and department B, respectively. Patients discharged from the hospital after surgery in one of the two departments can, if they want to, come to the follow-up department for further guidance and examinations. We are going to assume that the number of times a patient who have had surgery in department A is visiting the follow-up department is Poisson distributed with mean value $\mu_A = 1.4$. Correspondingly, we assume that the number of times a patient who have had surgery in department B is visting the follow-up department is Poisson distributed with mean value $\mu_B = 0.81$.

We are also going to consider it as known that of all patients having surgery in the two departments, 66% have surgery in department A.

a) Compute the probability that a patient from the department A have no visits to the follow-up department.

Given that a patient from department A has at least one visit to the follow-up department, find the probability that he or she has more than two visits to the follow-up department.

Compute also the probability that a random patient has no visits to the follow-up department.

Assume that one year department A makes surgery in $n_A = 16\ 302$ patients and that the same year department B makes surgery in $n_B = 8\ 398$ patients. Let X denote the number of these $n_A + n_B$ patients that have no visits to the follow-up department, and let Y denote the total number of visits these patients have to the follow-up department. Assume moreover that patients behave independently of each other.

b) Compute the mean and standard deviation of X?

What type of distribution has Y? Remember to give reason for your answer.

Component i	1	2	3	4	5	6	7	8	9	10
z_i	1.0	1.5	4.0	1.0	2.0	2.5	3.0	2.0	2.0	4.5
y_i	707	217	326	292	477	285	234	243	204	260

Table 1: Stress factors and corresponding observed life times. With these numbers we get in particular $\sum_{i=1}^{10}(y_i z_i)^2 = 6\ 736\ 616$.

Compute the mean and standard deviation of Y?

Problem 3 Life time for a new type of mechanical components

The life time Y (in number of days) for a new type of mechanical components is to be studied. It is known that the life time among other things depends on the temperature and pressure where the components are to be used. The effect of these conditions is quantified in a so called *stress factor*, z. Based on experience from similar components we assume the density function for Y for a given stress factor z to be given by

$$f(y) = \begin{cases} \frac{2yz^2}{\theta^2} \cdot e^{-\frac{y^2 z^2}{\theta^2}} & \text{for } y > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where $\theta > 0$ is a parameter describing the quality of the components. Throughout this problem we assume the life time for different mechanical components to be independent of each other.

a) Assume in this item that $\theta = 1000$ and $z = 1$.

Find the cumulative distribution function for Y, $F(y)$.

Find the probability that the life time Y is larger than 500 days.

Find the median for Y.

To study the quality of the new type of components one has observed the life time for $n = 10$ such components. Let $z_1, z_2, \ldots, z_n$ denote the stress factors under which these components where used and let $Y_1, Y_2, \ldots, Y_n$ denote the corresponding life times. The observed values are given in Table 1.

b) Show that the maximum likelihood estimator (MLE) for θ is

$$\widehat{\theta} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i z_i)^2}.$$

Compute the estimate.

For $i = 1, 2, \ldots, n$ let $U_i = \frac{2Y_i^2 z_i^2}{\theta^2}$.

c) Use the transformation formula to find the density function of U_i and use this to show that U_i is χ^2 distributed with 2 degrees of freedom.

Thereafter use this to argue that

$$\frac{2n\widehat{\theta}^2}{\theta^2} \sim \chi^2_{2n}.$$

d) From the result in **c**), derive a $(1-\alpha)\cdot 100\%$ confidence interval for θ.

Compute the numerical interval for the data in Table 1 when $\alpha = 0.05$.

Problem 4 Infection after surgery

A couple of years ago a surgery department underwent a reorganisation in order to reduce the risk of patients getting an infection.

Let p_F denote the probability that a random patient would get an infection before the reorganisation, and let p_E denote the corresponding probability after the reorganisation. For a period of one year before the reorganisation let n_F and X_F denote the number of patients in the surgery department and the number of these that got an infection, and let let n_E and X_E denote the corresponding quantities for a period of one year after the reorganisation. We assume different patients to get infections independent of each other.

In the following we use $\widehat{p}_F = \frac{X_F}{n_F}$ as an estimator for p_F, and correspondingly $\widehat{p}_E = \frac{X_E}{n_E}$ as an estimator for p_E.

a) Formulate the central limit theorem.

Assuming both n_F and n_E to be large, explain how the central limit theorem gives that $\widehat{p}_E - \widehat{p}_F$ is approximately normally distributed. (*Hint: explain first why $\widehat{p}_E$ and $\widehat{p}_F$ are approximately normally distributed.*)

Show that

$$\mathrm{E}\left[\widehat{p}_E - \widehat{p}_F\right] = p_E - p_F \quad \text{and} \quad \mathrm{Var}\left[\widehat{p}_E - \widehat{p}_F\right] = \frac{p_E(1-p_E)}{n_E} + \frac{p_F(1-p_F)}{n_F}.$$

	Number of patients	Patients who got an infection
Before	2 021	186
After	1 919	135

Table 2: The number of patients and the number of these patients who got an infection in some periods of one year before and after the reorganisation of the surgery department.

We now want to formulate a hypothesis test to test whether the reorganisation of the surgery department has been successful.

b) Formulate the null and alternative hypotheses for this situation. Specify what test statistic you want to use and discuss what the distribution of your test statistic is (approximately) when the null hypothesis is true.

Find the p-value of the test when the data is as specified in Table 2. Discuss whether or not you would conclude that the reorganisation has been successful.

Assuming the surgery department in the coming year will continue with the new organisation we want to find a prediction interval for the number of their patients that in the coming year will get an infection.

c) Assuming the number of patients in the coming year will be $m = 2\ 000$, derive a formula for a $(1 - \alpha) \cdot 100\%$ prediction interval for the number of these m patients that will get an infection related to their operation.

Using the values in Table 2 and $\alpha = 0.10$, calculate the prediction interval numerically.