LINKÖPINGS UNIVERSITET
Matematiska institutionen
Matematisk statistik

Kurskod: TAMS24
Provkod: TEN1
Datum: 2018-10-42
Institution: MAI

Example exam in Statistics

TAMS24/TEN1 2018-10-42 8–12

You are permitted to bring:

- a calculator (no computer);
- Formel- och tabellsamling i matematisk statistik (from MAI);
- Formel- och tabellsamling i matematisk statistik, TAMS65;
- TAMS24: Notations and Formulas (by Xiangfeng Yang).

Grading: 8-11 points giving grade 3; 11.5-14.5 points giving grade 4; 15-18 points giving grade 5. Your solutions need to be complete, well motivated, carefully written and concluded by a clear answer. Be careful to show what is random and what is not. Assumptions you make need to be explicit. The exercises are in number order.

Solutions can be found on the homepage a couple of hours after the finished exam.

1. A study of use of cannabis amongst adolescents (445 people) and their parents use of alcohol and/or other narcotics gave the following table.

		Usage (adolescents)		
		Never	Sometimes	Regularly
Usage (parents)	None	141	54	40
	One	68	44	51
	Both	17	11	19

   Using a suitable test, examine if there is a connection between the parents use of alcohol and/or narcotics and the adolescents' use of cannabis. (2p)

2. A company has stores in 4 large cities and measure sales during 8 weeks. They divide the raw numbers by the population in each city to normalize and the result (rounded off) can be seen in the table below.

		Week									
		1	2	3	4	5	6	7	8	$\overline{x}$	s
City	1	19	14	12	17	21	16	15	19	16.50	2.88
	2	8	12	9	9	8	10	11	8	9.38	1.51
	3	13	6	6	8	11	8	9	9	8.75	2.38
	4	17	18	17	22	16	18	14	21	17.88	2.59

We assume that the data are observations of independent random variables such that on row i, we have observations of $N(\mu_i, \sigma^2)$ (the variance is σ^2). We assume that all rows have the same variance. Construct two-sided confidence intervals for $\mu_i - \mu_j$ such that the *simultaneous* degree of confidence is $\geq 94\%$. Are there differences between the cities? (3p)

3. For a number of coal driven power plants with similar construction, one has measured the level y of sulphur dioxide release (ppm) and the effect x (GW). A helpful technician has provided the following computer analysis of the data. The model used was

$$Y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon,$$

where $\epsilon \sim N(0, \sigma^2)$ (the variance is σ^2) and different measurements are assumed to be independent. The regression line obtained was $y = 204 - 638x + 959x^2$ and we have the following data:

i	$\widehat{\beta}_i$	$d(\widehat{\beta}_i)$
0	204.46	82.95
1	-638.0	298.5
2	959.2	263.6

Analysis of variance

	Degrees of freedom	Square sum
REGR	2	15049.3
RES	6	254.3
TOT	8	15303.6

$$(X^T X)^{-1} = \begin{pmatrix} 162.38 & -581.90 & 508.70 \\ -581.90 & 2102.17 & -1850.58 \\ 508.70 & -1850.58 & 1639.75 \end{pmatrix}$$

(a) How many power plants were examined? (1p)

(b) Is the second degree term necessary? Motivate your answer at the level 5%. (1p)

(c) Find a 95% confidence interval for $\mathrm{E}(Y)$ when the effect is 0.5 GW. (2p)

4. A company is measuring the quality of work and for a certain variable used 25 samples $x_1, \ldots, x_{25}$ are collected independently during a day. It is reasonable to assume that this sample comes from the distribution $N(\mu,\ \sigma^2 = 1.2^2)$ (where $\sigma^2 = 1.2^2$ is the variance). The company wants to have $\mu > 30$ and needs help performing the test $H_0 : \mu = 30$ against $H_1 : \mu > 30$.

(a) If $\overline{x} = 30.35$, carry out the test at the significance level 5%. (1.5p)

(b) For which values of μ is the power of the test at least 75%? (1.5p)

5. The life time of a certain type of electrical components has the probability density

$$f(x) = a^2 x e^{-ax}, \quad x \geq 0,$$

where $a > 0$ is an unknown constant. During a test-run, 50 such components were found to have a cumulative life time of 250 time units (the 50 life times added together).

(a) Find a reasonable point estimate for a. (2p)

(b) Find a 95% confidence interval – one-sided bounded from below –for the expected life time of such a component. (2p)

6. Prove that S^2 is an unbiased estimate for σ^2. (2p)