LINKÖPINGS UNIVERSITET
Matematiska institutionen
Matematisk statistik

Kurskod: TAMS24
Provkod: TEN1
Datum: 2019-01-09
Institution: MAI

Exam in Statistics

TAMS24/TEN1 2019-01-09

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 (sufficient limits): 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 reasonable question is which one of Morbid Angel's first bunch of studio albums is the best (so no live albums and to avoid confusion not the Abominations of Desolation album[1]). The following data was collected from two sources on the internet.

Album Title	Web page	
	Nuclear War Now!	Metalstorm.net
Altars of Madness	67	82
Blessed Are The Sick	18	34
Covenant	11	32
Domination	1	21
Formulas Fatal To The Flesh	6	4
Gateways to Annihilation	1	10
Heretic	0	2

Use a suitable test with significance level 0.01 to see if there is a difference between opinions on the two sites. (2p)

2. Lina is experimenting with water cooling for her computers. She's measured temperatures of the cpu:s in 5 different computers (those who survived the experiment), first with conventional cooling and then after switching to water cooling.

[1] Since it's obviously the best.

	Temperature				
Computer:	C-1	C-2	C-3	C-4	C-5
Conventional cooling	55	36	55	64	53
Water cooling	50	38	39	50	44

We assume that the temperatures are normally distributed and that different computers are independent.

(a) Find 95% confidence intervals for the expected temperatures with conventional cooling and water cooling, respectively. (2p)

(b) Perform a test for if there is a difference in the expected temperature between the cooling techniques at the level 5%. (2p)

(c) Find a confidence interval $I_{\sigma^2} = [0, a)$ for the variance of the temperature using water cooling. Use the degree of confidence 90%. (1p)

3. In statistics, one frequently works with *stochastic processes*. One such example could be expressed as $X(n)$ for $n = 1, 2, 3, \ldots$, where $X(n)$ is a random variable for each n. Let one such process $X(n)$ satisfy the following. It has expectation 0 for every n (that is, $\mathrm{E}(X(n)) = 0$) and if $Y(n)$ is the random vector $Y(n) = (\, X(n),\ X(n-1),\ X(n-2)\,)^T$, then the covariance matrix is given by

$$C_{Y(n)} = \mathrm{E}(Y(n)Y(n)^T) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

for every $n = 3, 4, 5, \ldots$.

Find a linear predictor $\widehat{X}(n) = aX(n-1) + bX(n-2)$ of $X(n)$ that minimizes the quadratic error. In other words, find a and b such that $\mathrm{E}((\widehat{X}(n) - X(n))^2)$ is minimal. (2p)

4. Crawford Tillinghast has built a machine that enables people to see and interact with alternate dimensions. It works by stimulating the pineal gland in the brain by means of resonance waves. While building his machine, he measured the frequency of the waves as a function of the voltage he applied, and he also took note of if the measurement was made at night (represented by 0) or during the day (represented by 1).

Frequency (f)	1.86	2.41	3.26	3.88	4.64	5.50	6.40
Voltage (v)	1	2	3	4	5	6	7
Day/Night (u)	0	0	1	0	0	1	1

Crawford believes that the frequency depends linearly on the voltage, but he isn't sure that the time of day is important (but he gets his most spectacular results at night). He considers the following two models:

$$\text{Model 1:} \quad F = \beta_0 + \beta_1 v + \epsilon$$

and

$$\text{Model 2:} \quad F = \beta_0 + \beta_1 v + \beta_2 u + \epsilon,$$

where $\epsilon \sim N(0, \sigma^2)$ and different measurements are assumed to be independent. The following calculations has already been carried out.

Model 1:

i	$\widehat{\beta}_i$	$d(\widehat{\beta}_i)$
0	0.9671	0.0984
1	0.7564	0.0220

Analysis of variance

	Degrees of freedom	Square sum
REGR	1	16.0212
RES	5	0.0678
TOT	6	16.0889

Model 2:

i	$\widehat{\beta}_i$	$d(\widehat{\beta}_i)$
0	0.9866	0.0923
1	0.7370	0.0250
2	0.1363	0.1009

Analysis of variance

	Degrees of freedom	Square sum
REGR	2	16.0424
RES	4	0.0466
TOT	6	16.0889

(a) Is the term in model 2 corresponding to day/night meaningful? Carry out a test at the 1%-level. What is the interpretation of your result? (2p)

(b) Find a 95% confidence interval for β_1 using model 2. (1p)

5. In a game of *Death Adder Roulette*, played out in the Australian outback, people take turns in trying to pet a venomous snake (traditionally a death adder) on the head. The game is played until someone is bitten. Assume that the probability of being bitten is constant.

(a) Assume that a person is bitten at try number n. Find a reasonable (using n) point estimate for p and calculate the expectation for the estimator. Is the estimator unbiased? (2p)

(b) One participant claims that the current snake i feisty so that $p = 0.4$. In one game, the first person was bitten at the fifth try. Test the hypothesis $H_0 : p = 0.4$ versus $H_1 : p < 0.4$ at the significance level 5%. (1p)

(c) What is the power of the test at $p = 0.2$? (1p)

6. Suppose that $\boldsymbol{Y} = (Y_1\ Y_2\ \cdots\ Y_k)^T \sim N(\boldsymbol{0}, I_k)$, where I_k is the $k \times k$ identity matrix. If $A \in \mathbf{R}^{k\times k}$ is such that A is symmetric, the column rank of A is l $(0 < l \leq k)$, and $A^2 = A$, derive the distribution for $\boldsymbol{Y}^T A\boldsymbol{Y}$. (2p)