Chalmers tekniska högskola - Göteborgs universitet **Exam: 2021-10-26** $8^{30} - 12^{30}$.
MVE055/MSG810 Matematisk Statistik och Diskret Matematik

Teacher on call: Moritz Schauer, phone +46 31 772 3029.

Permitted materials: Beta-booklet, attached formula sheet and Chalmers-approved calculator. The correct, well motivated solution gives the points given in parentheses for each task. Grade limits: Chalmers: 3: 12-17p, 4: 18-23p, 5: 24-30p, GU: G: 12-21, VG: 22-30.

---

1. Let X be a random variable with binomial distribution with $n = 10$ and $\frac{1}{2} \le p < 1$ unknown. Alice wants to know if $p > \frac{1}{2}$.

   (a) Formulate a null-hypothesis and an alternative hypothesis. (1p)

   $$H_0\colon p = \frac{1}{2} \quad H_1\colon p > \frac{1}{2}$$

   (b) As a decision rule, Alice proposes to reject the null hypothesis if $X > 7$. Determine the error of the first kind (type 1 error). (2p)

   To calculate the error, assume H_0 so $p = \frac{1}{2}$. Then

   $$\alpha = P(X \in \{8, 9, 10\}) = \left(\binom{10}{8} + \binom{10}{9} + \binom{10}{10}\right)\frac{1}{2^{10}} = \frac{45 + 10 + 1}{2^{10}} = \frac{7}{128} = 0.0546$$

   (c) Bob prefers to perform the hypothesis test at significance level $\alpha = 0.02$. Suggest an appropriate modification of Alice' test rule. (2p)

   Alice test has level $\alpha > 0.05$. But Bob can use the following rule: reject the null hypothesis if $X > 8$, with $\alpha = 0.0107$.

2. The ping time (latency) is the time it takes for a small data set (packet) to be transmitted from a device to a server on the internet and back to the device again. In online play, a small ping time is important.

   (a) The ping time X of your device is exponentially distributed with mean $\mathrm{E}(X) = 20$ ms. Find the probability that a packet takes longer than 60 ms. (1p)

   Choose the parameter such that the *mean* equals 20. Then

   $$P(X > 60) = 1 - F_X(60) = \exp(-60/20) = 0.0497...$$

   (b) Find the probability that a packet takes longer than 120 ms given that it already took longer than 60 ms. (2p)

   $$\frac{P(X > 120)}{P(X > 60)} = \frac{\exp(-120/20)}{\exp(-60/20)} = \exp(-60/20) = 0.0497...$$

   (c) Later that day you have connectivity issues. Your investigation shows that your packets are lost with 1 % probability. If a packet is not lost, it will return after the exponentially distributed time with mean 20 ms. Find the probability that the package will never return if it did not return yet after 60 ms. (3p)

   $$\frac{P(\text{ lost})}{P(X>60|\text{not lost})P(\text{not lost}) + P(X>60|\text{lost})P(\text{ lost})} = \frac{0.01}{\exp(-60/20)0.99 + 1\cdot 0.01} = 0.1689...$$

3. Neural networks are trained using an iterative optimization process that requires a loss function to calculate the model error. A simple statistical model for the change $x_{i+1} - x_i$ of the loss x_i during the training iteration i is

$$x_{i+1} - x_i = a(\mu - x_i) + \epsilon_i, \quad i = 1, 2, \ldots,$$

where $\epsilon_i \sim N(0, \sigma^2)$ are independent error terms. Here the loss x_i after iteration i is recorded during the training process and μ and a are unknown constants of interest: μ is the training error in the fully trained model and $a > 0$ is the learning rate.

The table below gives the training loss x_i for a particular training run of such a neural network for $i = 1, \ldots, 10$ and also $i = 11$.

i	1	2	3	4	5	6	7	8	9	10	(11)
x_i	0.495	0.193	0.085	0.042	0.017	0.016	0.007	0.019	0.013	0.011	(0.009)

With $y_i = x_{i+1} - x_i$, $i = 1, \ldots, 10$ (the change in loss) and coefficients $\beta_0 = a\mu$ and $\beta_1 = -a$ we can rewrite the model as linear regression model

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i \in \{1, \ldots, 10\}.$$

$$\sum_{i=1}^{10} x_i = 0.8980 \quad \sum_{i=1}^{10} y_i = -0.4860$$

$$\sum_{i=1}^{10} x_i^2 = 0.29251 \quad \sum_{i=1}^{10} y_i^2 = 0.10561 \quad \sum_{i=1}^{10} x_i y_i = -0.17528$$

Remark: Calculate with sufficient precision.

(a) Find the regression line for the change in loss y_i with current loss x_i as explanatory variable. (2p)

$$y_i = 0.0072 - 0.6213 x_i + \epsilon_i$$

(b) Use this to find the training error in the fully trained model μ and the learning rate $a > 0$. Do you expect substantially smaller loss if this model (which was trained for 11 iterations) is trained for more iterations? (2p)

$$\mu = 0.0115, \quad a = 0.621.$$

The model seems to be fully trained, $x_{11} \approx \mu$.

(c) Is β_0 significantly larger than 0? Formulate hypotheses and test at significance level $\alpha = 0.05$. Assume normal errors with unknown variances σ^2. (2p)

The question suggests a one-sided test. $Q_0 = 0.00020$ and therefore $S = 0.0050$. $t_c = 1.86$ is the $1 - \alpha$ quantile of the $t(10 - 2)$-distribution.

$$T = \frac{\beta_0 - 0}{S\sqrt{1/10 + \bar{x}^2/S_{xx}}} = 3.812 \quad \text{where e.g. } S_{xx} = \left(10 \cdot 0.29251 - (0.898)^2\right)/10$$

is more extreme than t_c, hence reject the null hypothesis that the intercept is 0 and conclude that it is larger.

(d) Find a 90 % prediction interval for x_{12} given $x_{11} = 0.009$. (Hint: Compute a prediction interval for y_{12} first.) (2p)

$t_c = 1.86$ is the 0.95-quantile of the $t(10-2)$-distribution. The formula for a prediction interval for y gives

$$I_y = (\beta_0 + \beta_1 x_{11}) \pm tS\sqrt{1 + 1/n + (x_{11} - \bar{x})^2/S_{xx}}$$
$$= [-0.00844, 0.0116].$$

That implies

$$I_{x_{12}} = [x_{11} - 0.00844, x_{11} + 0.0116] = [0.00056, 0.0207].$$

4. A student has class on Tuesday mornings at 10:00. She can choose between the early 9:20 bus and the late 9:50 bus.

If she chooses the 9:20 bus, she will be in time. The next week she will chose the early bus or the late bus with equal probability.

But if she chooses the 9:50 bus, she will be late with probability 1/3. If she is late, she has a bad conscience and takes the 9:20 bus the next week. If she is not late, she will take the 9:50 bus again the next week.

(a) The behaviour can be modelled by a Markov chain with two states: taking bus 1 or taking bus 2. Write down the transition matrix of the chain, (1p)

$$\mathbf{P} = \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}.$$

$$\mathbf{P} = \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix}.$$

(b) Find the stationary distribution $q = \begin{bmatrix} q_1 & q_2 \end{bmatrix}$ by solving the equation system $q\mathbf{P} = q$. (2p)

$$q_1 = q_1/2 + q_2/3 \quad \text{and} \quad q_2 = q_1/2 + 2q_2/3 \quad \text{and} q_1 + q_2 = 1$$

give

$$q = \begin{bmatrix} 2/5 & 3/5 \end{bmatrix}$$

(c) Is q_i equal to the probability of the Markov chain being in the state $i = 1, 2$ in the long run is this case? (What condition does $\mathbf{P}$ need to satisfy? Does it?) (2p)

$\mathbf{P}$ is regular. Therefore yes.

(d) On which proportion of Tuesdays will the student be late in the long run? (2p)

She takes the late bus with probability 3/5 in the long run, and *if*, she is late with probability 1/3, so she will be late a proportion of $3/5 \cdot 1/3 = 1/5$ of Tuesdays.

5. Each number in the Fibonacci sequence is the sum of the two preceding ones, starting from $a_0 = 0$ and $a_1 = 1$.

(a) Use generating functions to find an explicit form for a_n, (4p)

$$\begin{cases} a_0 = 0 \\ a_1 = 1 \\ a_2 = 1 \\ a_3 = 2 \\ a_n = a_{n-1} + a_{n-2} \quad n > 1 \end{cases}$$

Hint:

$$\frac{x}{1-x-x^2} = \frac{1}{\sqrt{5}}\left(\frac{1}{1-\varphi x} - \frac{1}{1-\psi x}\right)$$

where $\varphi = \frac{1+\sqrt{5}}{2} = 1.618\ldots$ is the golden ratio and $\psi = \frac{1-\sqrt{5}}{2} = 0.618\ldots$.

We are looking for $a_0, \ldots, a_n, \ldots$ Let

$$f(x) = \sum_{k=0}^{\infty} a_k x^k.$$

Now

$$f(x) = a_0 + a_1 x + \sum_{k=1}^{\infty}(a_{k-1} + a_{k-2})x^k = 0 + x + (x + x^2)f(x)$$

$$f(x) = \frac{x}{1-x-x^2} = \frac{1}{\sqrt{5}}\left(\frac{1}{1-\varphi x} - \frac{1}{1-\psi x}\right) = \frac{1}{\sqrt{5}}\left(\sum_{k=0}^{\infty}(\varphi x)^k - \sum_{k=0}^{\infty}(\psi x)^k\right)$$

using the hint and the geometric series.
Thus we have identified the expansion of f and read its coefficients as

$$a_n = \frac{\varphi^n - \psi^n}{\sqrt{5}}.$$

(b) Also prove the hint. (2p)

$$\frac{1}{\sqrt{5}}\left(\frac{1}{1-\varphi x} - \frac{1}{1-\psi x}\right) = \frac{1}{\sqrt{5}}\left(\frac{-\psi x + \varphi x}{(1-\varphi x)(1-\psi x)}\right) =$$

$$= \frac{1}{\sqrt{5}}\left(\frac{-\psi x + \varphi x}{1 - \varphi x - \psi x + \varphi\psi x^2}\right) = \frac{1}{\sqrt{5}}\left(\frac{\sqrt{5}x}{1-x-x^2}\right)$$

Lycka till!