Chalmers tekniska högskola - Göteborgs universitet **Exam: 2022-10-25** $8^{30} - 12^{30}$.
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. You need a fair coin to decide who does the dishes, you or your friend John von Neumann. You only have two unfair coins (one from copper, one from silver). You can assume that each coin has probability of heads $p = 2/3$.

   (a) You throw your two coins a single time. What is the probability that both coins show the same symbol? Assume independence. (2p)

   P(the same) = P(both head) + P(both tails) $= p^2 + (1-p)^2 (= 5/9)$.

   (b) Your friend proposes the following: "We throw both of them. When both coins show the same thing, we throw them both again. We do this until the coins show a different side each. Then we look at the copper coin. If it shows heads, I'll do the dishes, otherwise you."
   Is that fair? Compute the conditional probability that the copper coin shows heads given that both coins show a different result. (2p)

   P(copper coin heads ∩ both coins different)/P(both coins different) = (P(copper coin heads ∩ copper coin head, silver coin tails) + P(copper coin head ∩ copper coin tails, silver coin heads)) /P(both coins different) $= \frac{1/3 \cdot 2/3 + 0}{4/9} = 1/2$.

   (c) On average, how many times do you have to throw the pair of coins, until you can read off the results (they need to show a different face for that)? Model the number of times k you need to throw both coins to get the result. Choose the right distribution and report its name, the parameter and the mean. (To get the result you have to throw at least one time, of course, so $k \geq 1$.) (2p)

   The distribution is geometric with parameter (success probability in single experiment) $1 - (p^2 + (1-p)^2) = 2p - 2p^2$
   The expectation is
   $$1/(2p - 2p^2)(= 9/4).$$

2. In a sample of $n = 10$ human subjects height is measured. The table below gives the measurements y_i in m and x_i is equal to 1 if the subject is biologically male and zero otherwise.

x_i	1	0	1	0	0	0	1	0	1	1
y_i	1.82	1.77	1.72	1.63	1.59	1.58	1.87	1.67	1.77	1.98

   (a) Find the regression line of the height y_i with the variable x_i as explanatory variable. What is the average male height according to regression line/the model? (2p)

   $$y = 1.648 + 0.184x$$

(b) Is there any significant linear relation between the height score and the variable x encoding biological sex? Formulate a null hypothesis and test at significance level $\alpha = 0.05$. Assume normal errors with unknown variance σ^2. (2p)

$Q_0 = 0.06356$, $s^2 = 0.00794 = 0.08913^2$, $df = 8$

$$t^\star = 2.3060, |T| = 3.2639, p = 0.01146$$

Yes, there is a significant relationship.

(c) Find 95% confidence intervals for the average male ($x = 1$) and female ($x = 0$) height. (2p)

CI female height $[1.56, 1.74]$. CI male height $[1.741.92]$.

(d) Find 95% predictive intervals for the male ($x = 1$) and female ($x = 0$) height. (2p)

Prediction interval female height $[1.42, 1.87]$, Prediction interval male height $[1.61, 2.06]$.

3. X is a continuous random variable with probability density

$$f_X(x) = \begin{cases} 0 & x \leq 0 \\ c \cdot x & x \in (0, \frac{1}{2}] \\ c \cdot (1-x) & x \in (\frac{1}{2}, 1] \\ 0 & x \geq 1 \end{cases}.$$

(a) Find the constant $c > 0$ (the value of c that makes f_X into a probability density.) (2p)

$$\int f_X(x) = c \cdot \left[\frac{1}{2}x^2\right]_0^{\frac{1}{2}} + c \cdot \left[x - \frac{1}{2}x^2\right]_{\frac{1}{2}}^1 = 1 \Rightarrow c = 4$$

(b) Find the distribution function of X. (2p)

$$F_X(x) = \begin{cases} 0 & x \leq 0 \\ 2 \cdot x^2 & x \in (0, \frac{1}{2}] \\ 1 - 2 \cdot (1-x)^2 & x \in (\frac{1}{2}, 1] \\ 1 & x \geq 1 \end{cases}.$$

(c) Let $Y = 1 - X$. Find either the distribution function of Y or, alternatively, the density of Y.

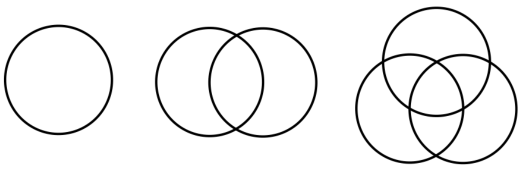
Answer: Which is correct?

(i) X and Y have the same distribution, but $P(X = Y) = 0$.
(ii) X and Y have the same distribution, but $P(X = Y) = \frac{1}{2}$.
(iii) X and Y have a different distribution.

(2p)

The density and the distribution functions of Y are the same as of X. Thus (i) is correct.

4. A Venn diagram divides the plane into a number of different areas. A single circle divides the plane into 2 regions, inside and outside. We have $a_1 = 2$, $a_2 = 4$, $a_3 = 8$ areas in Venn diagrams with $k = 1, 2$, and 3 circles, as shown below.



In a diagram with $k - 1$ circles, drawing another circle produces $2(k - 1)$ new areas. Therefore $a_4 = 14$ and in general

$$a_k = a_{k-1} + 2(k-1), \quad k \geq 2.$$

Show or disprove that the generating function of the sequence $\{a_k\}$ is given by

$$A(x) = \frac{2x^2}{(1-x)^3} + \frac{2x}{1-x}.$$

You don't need to find the formula for a_k (which is $a_k = (k-1)k + 2$.)

Hint:

$$\sum_{k=0}^{\infty}(k+1)x^k = \frac{1}{(1-x)^2}, \quad |x|<1.$$

(5p)

The recursion gives $\sum_{k\geq 2} a_k x^k = \sum_{k\geq 2}\left(a_{k-1}x^k + 2(k-1)x^k\right) = x\sum_{k\geq 2} a_{k-1}x^{k-1} + 2x\sum_{k\geq 2} 2(k-1)x^{k-1}$. Using

$$\sum_{k=1}^{\infty} k\cdot x^{k-1} = \frac{1}{(x-1)^2}$$

we have

$$x\sum_{k=2}^{\infty} 2(k-1)x^{k-1} = (2x^2)/(x-1)^2 \text{ when } |x|<1$$

which gives

$$-2x + \sum_{k\geq 1} a_k x^k = x\sum_{k\geq 1} a_k x^k + x(2x)/(x-1)^2.$$

Thus $A(x)$ solves $-2x + A = xA + x(2x)/(x-1)^2$ and this gives
$A = (2x + x(2x)/(x-1)^2)/(1-x)$. or

$$A = \frac{2x^2}{(1-x)^3} + \frac{2x}{1-x}$$

From there on can continue to show

$$a_k = (k-1)k + 2.$$

5. You are selling Chalmers approved calculators to your fellow students during working hours. On the evening before the first day of business you buy three 3 of them in the student shop. On any day you sell each of the calculators you have in the morning independently with probability 1/2. The next day you start with the remaining number of calculators, but if you have no calculators left in the evening because you sold them all during the day, you get 3 new ones for the next day. Thus in the morning you have either 1, 2, or 3 calculators in your inventory.

   Model the number of calculators in the inventory in the morning by a Markov chain with three states E_1, E_2, E_3 corresponding to having 1, 2, or 3 calculators in the morning. The chain is started on the first morning with three calculators in the inventory.

   (a) What is the probability to start on the second morning with 3 calculators? (Think of the different ways this could happen.) (1p)

   Either sell all of them (with probability $1/2^3$,) or sell none of them (also with probability $1/2^3$.) That makes 2/8.

   (b) Find the transition matrix. Hint: The probability to sell 1 calculator if you have 3 is $\binom{3}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^{3-1} = 3/8$. (2p)

$$\mathbf{P} = \begin{array}{c} \begin{array}{ccc} E_3 & E_2 & E_1 \end{array} \\ \begin{bmatrix} ? & 3/8 & ? \\ ? & ? & ? \\ ? & 0 & ? \end{bmatrix} \end{array} \begin{array}{c} \\ E_3 \\ E_2 \\ E_1 \end{array}$$

$$\begin{array}{cc} & \begin{array}{ccc} E_3 & E_2 & E_1 \end{array} \\ \mathbf{P} = & \begin{bmatrix} 2/8 & 3/8 & 3/8 \\ 1/4 & 1/4 & 1/2 \\ 1/2 & 0 & 1/2 \end{bmatrix} \begin{array}{c} E_3 \\ E_2 \\ E_1 \end{array} \end{array}$$

(c) Check if $q = \begin{bmatrix} 4/11 & 2/11 & 5/11 \end{bmatrix}$ is the stationary distribution. (2p)

Verify

$$\begin{bmatrix} 4/11 & 2/11 & 5/11 \end{bmatrix} \mathbf{P} = \begin{bmatrix} 4/11 & 2/11 & 5/11 \end{bmatrix}.$$

Lycka till!