FINAL EXAM,
FIRST COURSE IN PROBABILITY AND STATISTICS

- **Time**: 20.2.2019, 9:00-12:00
- **Equipment**: Calculator and one sheet (A4) of hand-written notes, written on one side only.
- Answer each problem on a separate page. Each problem is worth 6 points.
- Motivate all solutions carefully. Answers without motivation give no points.
- Mark your course code on the front page.

Problem 1

A red, white and blue die are rolled (all three dice are fair and six-sided). Denote their outcomes respectively by A (red die), B (white die) and C (blue die).

(a) Compute the conditional probability $P(A < C|A = i)$, $i = 1, \dots, 6$. (1p)
(b) Compute the probability $P(A < C)$. (1p)
(c) Compute the conditional probability

$$P(\{A < B\} \cap \{A < C\}|A = i), \qquad i = 1, \dots, 6. \qquad (1p)$$

(d) Compute the joint probability $P(\{A < B\} \cap \{A < C\})$. (1p)
(e) Compute the conditional probability $P(A < B|A < C)$. (2p)

Problem 2

60% of the Finnish population is "young", by which we mean below 50 years, and the rest is "old". 35% of young people use glasses, whereas 85% of old people do.

A sample of 100 individuals are selected at random (with replacement). Let X be the number of young people in the sample, and let Y be the number of people in the sample who wear glasses.

(a) Compute $E(X)$. (1p)
(b) Compute $E(Y)$. (2p)
(c) Compute the covariance $\mathrm{Cov}(X, Y)$. (3p)
(hint: write X and Y as sums of indicator variables.)

Problem 3

100 random numbers are drawn independently from the continuous uniform distribution on $[-1, 2]$. Let X be the number of positive numbers drawn. Use the normal approximation to estimate $P(X < 60)$. (6p)

Problem 4

The time X (in seconds) from when I leave my office until I jump on the metro can be modelled as a constant time c (to walk to the metro station) plus an exponentially distributed time with rate λ (waiting). So the probability density function of X is

$$f(t) = \begin{cases} \lambda e^{-\lambda(t-c)}, & t \geq c \\ 0, & \text{otherwise} \end{cases}$$

The waiting times on different days are supposed to be independent. The last five days, X was $185, 400, 250, 500, 375$.

(a) Write down the likelihood function for the unknown parameters c and λ. (2p)
(b) Compute the maximum likelihood estimate of c. (2p)
(c) Compute the maximum likelihood estimate of λ. (2p)

1. Statistical tables

Kertymäfunktion $\Phi(z)$ arvoja / Values of the cumulative distribution function $\Phi(z)$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

1:

a) $\mathbb{P}(A<C \mid A=i) = \mathbb{P}(i<C) = \frac{6-i}{6} = 1-\frac{i}{6}$ 1p

b) $\mathbb{P}(A<C) = \sum_{i=1}^{6} \mathbb{P}(A=i)\mathbb{P}(A<C|A=i)$

$= \sum_{i=1}^{6} \frac{1}{6}\left(1-\frac{i}{6}\right)$

$= \frac{1}{6}\left(\frac{5}{6}+\frac{4}{6}+\frac{3}{6}+\frac{2}{6}+\frac{1}{6}\right) = \frac{15}{36}$ 1p

c) $\mathbb{P}(A<C \,\&\, A<B \mid A=i) = \mathbb{P}(i<C \,\&\, i<B)$

$\underset{\text{indep}}{=} \mathbb{P}(i<C)\mathbb{P}(i<B) = \left(1-\frac{i}{6}\right)^2 = \frac{(6-i)^2}{36}$ 1p

d) $\mathbb{P}(A<C \,\&\, A<B) = \sum_{i=1}^{6} \mathbb{P}(A=i)\,\mathbb{P}(A<C \,\&\, A<B \mid A=i)$

$= \sum_{i=1}^{6} \frac{1}{6}\cdot\frac{(6-i)^2}{36}$

$= \frac{1}{216}(25+16+9+4+1) = \frac{55}{216}$ 1p

e) $\mathbb{P}(A<B \mid A<C) = \frac{\mathbb{P}(A<B \,\&\, A<C)}{\mathbb{P}(A<C)}$ 1p

$= \frac{55/216}{15/36} = \frac{11}{18}$ 1p

2

a) $\mathbb{E}[X] = 100 \cdot \mathbb{P}(\text{young}) = 100 \cdot 0{,}6 = 60.$

(1p)

b) $\mathbb{E}[Y] = 100 \cdot \mathbb{P}(\text{glasses})$

$$= 100\Big[\mathbb{P}(\text{young})\,\mathbb{P}(\text{glasses}|\text{young}) + \mathbb{P}(\text{old})\,\mathbb{P}(\text{glasses}|\text{old})\Big]$$

$$= 100(0{,}6 \cdot 0{,}35 + 0{,}4 \cdot 0{,}85) = 55$$

(2p)

c) Let $X_i = \begin{cases} 1 & \text{if } i^{th} \text{ individual young} \\ 0 & \text{otherwise} \end{cases}$

$Y_i = \begin{cases} 1 & \text{if } i^{th} \text{ individual has glasses} \\ 0 & \text{otherwise.} \end{cases}$

$$X = \sum_{i=1}^{100} X_i, \quad Y = \sum_{i=1}^{100} Y_i$$ (1p)

$$\mathrm{Cov}(X,Y) = \mathrm{Cov}\Big(\sum_i X_i, \sum_j Y_j\Big)$$

$$= \sum_{i,j} \mathrm{Cov}(X_i, Y_j)$$

$$= \sum_{i=j} \mathrm{Cov}(X_i, Y_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, Y_j)$$

X_i, Y_j independent ~~samples~~ if $i \neq j$

$$= 100\, \mathrm{Cov}(X_1, Y_1)$$ (1p)

2 cont.

c) $\mathrm{Cov}(X_1, Y_1) = \mathbb{E}[X_1 Y_1] - \mathbb{E}[X_1]\mathbb{E}[Y_1]$

$= \mathbb{P}[\text{young} \,\&\, \text{glasses}] - \mathbb{P}[\text{young}]\,\mathbb{P}[\text{glasses}]$

$= \mathbb{P}[\text{young}]\left(\mathbb{P}[\text{glasses} \mid \text{young}] - \mathbb{P}[\text{glasses}]\right)$

$= 0.6(0.35 - 0.55) = -0.12$

so $\mathrm{Cov}(X, Y) = 100 \cdot \mathrm{Cov}(X_1, Y_1) = -12.$

1p

This problem can also be solved via directly computing $\mathbb{E}[XY] = \mathbb{E}[\Sigma X_i \Sigma Y_j]$

$= \sum_{i,j} \mathbb{E}[X_i Y_j].$

Also In such case, give 1p for ~~the observation~~ writing down the sum, 1p for observing $X_i \perp Y_j$ if $i \neq j$, and 1p for correct computation.

3

Every number drawn is positive with probability $2/3$

So $X \sim Bin(100, 2/3)$ 1p

$\mathbb{E}[X] = \frac{200}{3}$, $Var[X] = 100 \cdot \frac{2}{3} \cdot \frac{1}{3} = \frac{200}{9}$ 1p

So by normal approximation,

$$\frac{X - \frac{200}{3}}{\sqrt{200/9}} \underset{\text{approx}}{\sim} \mathcal{N}(0,1)$$ 1p

$$\mathbb{P}[X<60] \underset{\text{Continuity correction}}{=} \mathbb{P}[X \le 59.5] = \mathbb{P}\left[\frac{X - \frac{200}{3}}{\sqrt{200/9}} \le \frac{59.5 - \frac{200}{3}}{\sqrt{200/9}}\right]$$ 1p

$$\approx \Phi\left[\frac{59.5 - \frac{200}{3}}{\sqrt{200/9}}\right] \approx \Phi(-1.52)$$ 1p

$$\approx 0.0643$$ 1p

4 a)

$$L(c,\lambda) = f_{c,\lambda}(185) \cdot f_{c,\lambda}(400) \cdot f_{c,\lambda}(250) \cdot f_{c,\lambda}(500) \cdot f_{c,\lambda}(375)$$

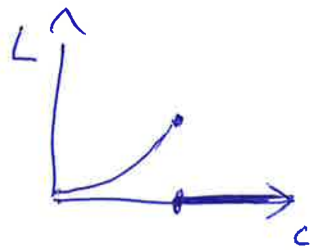
(1p)

$$= \begin{cases} \lambda^5 e^{-\lambda(185-c+400-c+250-c+500-c+375-c)} & \text{if } c \leq \min\{185, 400, 250, 500, 375\} \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \lambda^5 e^{-\lambda(1710-5c)} & \text{if } c \leq 185 \\ 0 & \text{otherwise} \end{cases}$$

(1p)

b) $L(c,\lambda)$ is positive and increasing in c for $c \leq 185$, so maximized by $c = 185$.



(2p)

4 c) $$\frac{dL}{d\lambda} = -\lambda^5(1710-5_c)\, e^{-\lambda(1710-5_c)} + 5\lambda^4 e^{-\lambda(1710-5_c)}$$

1p

(also ok to first take the logarithm and then differentiate)

L ~~maximized~~ extreme when

$$0 = \frac{dL}{d\lambda} \iff \lambda^5(1710-5_c) = 5\lambda^4$$

$$\iff \lambda = \frac{5}{1710-5_c} \overset{c=185}{=\!=} \frac{5}{785} = \frac{1}{157}.$$

Sign studies or second derivative show this extreme value is a maximum.

1p