MATH-232 Final Exam, Solutions

Exercise 1.

(a) *How many permutations of the letters of the word 'statistics' give a different sequence of letters?*

(b) *Among 200 students, what is the probability that at least two students have the same birthday (for example, August 14, without the year mattering) and the same eye color? Note: assume that the birthdays are independently and uniformly distributed among 365 days and that the eye colors are independently and uniformly distributed among blue, green, brown.*

Exercise 2. *Assume that a person infected by COVID transmits the virus to another person during a meeting with probability (i)* $1/2$ *if none wears a mask, (ii)* $1/10$ *if only one person wears a mask (irrespective of which), and (iii)* $1/100$ *if both wear masks. Assume that people wear masks independently with probability* $1/2$*. Given that Alice transmitted the virus to Bob during a meeting, what is the probability that Bob wore a mask during that meeting? Note: we do not know if Alice wore a mask.*

Exercise 3. *Let $X \sim \mathcal{N}(0,1)$ and $Y = X^2$.*

(a) *Are X and Y uncorrelated?*

(b) *Are X and Y independent?*

Exercise 4. *Let $X_1, \ldots, X_n$ be i.i.d. such that X_i equals 0 with probability 1/2 and π with probability 1/2, for all $i \in [n]$. Show that $\frac{1}{n}\sum_{i=1}^{n} X_i$ tends in probability to a scalar x, and find the value for x. Note: you cannot invoke theorems from the class, you need to prove the statement in full (possibly using the formulaire).*

Recall: Y_n tends to Y in probability if for any $\epsilon > 0$, $\mathbb{P}(|Y_n - Y| > \epsilon) \overset{n\to\infty}{\longrightarrow} 0$.

Exercise 5. *Let $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim \mathcal{N}_2\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\right)$. Can you find a 2×2 matrix A such that AX has i.i.d. $\mathcal{N}(0,1)$ components? Disprove or provide A.*

Exercise 6. *Let $Y_1, \ldots, Y_n$ be i.i.d. such that Y_i equals 1 with probability p and -1 with probability $1-p$, for all $i \in [n]$.*

(a) *Find an estimator of p using the method of moments.*

(b) *What is the MSE of your estimator in (a)?*

Recall: the MSE of an estimator $\hat{\theta}$ for a parameter θ is $\mathbb{E}(\hat{\theta} - \theta)^2$.

(c) *Find an estimator of p using the maximum likelihood method.*

Exercise 7. *Let $Y_1, \ldots, Y_n$ be i.i.d. $\mathcal{N}(\mu, 4)$ for some $\mu \in \mathbb{R}$.*

(a) *Find the values of a and b for which $(\sum_{i=1}^n Y_i/n - a)/b$ is distributed as $\mathcal{N}(0,1)$. Note: a, b may depend on μ, n.*

(b) *Using this pivot, give an equi-tailed two-sided confidence interval for μ with confidence level $\alpha = 0.05$ (i.e. $\alpha_L = \alpha_U = \alpha/2$) in the special case where $\sum_{i=1}^n Y_i/n = 1$ and $n = 100$. Note: the formulaire contains Gaussian quantiles.*

Exercise 8. *Consider the binary hypothesis test where a single random variable Y is observed. Under H_0, the random variable Y takes values $\{0, 1, 2\}$ with probabilities $\{1/4, 1/2, 1/4\}$ respectively, and under H_1, the random variable Y takes values $\{1, 2, 3\}$ with probabilities $\{1/4, 1/2, 1/4\}$ respectively.*

(a) *Give a test (i.e., define the set of values of Y for which you accept, or reject, H_0) such that the false negative probability is minimal under the constraint that the false positive probability is at most $1/3$. Note: the question is symmetrical if you exchange the role of the false positive and false negative probabilities, so it is ok if you flip the two definitions.*

(b) *Same question as in (a) under the constraint that the false positive probability is at most $1/5$.*