# Mathematical Probability (Math 6383)

## MIDTERM TEST. SOLUTIONS

1. Let $\Omega = [0,1]$ and let $\mathcal{A}$ consist of all subsets of $[0,1]$ that are either countable or have countable complements. Define $\mathbf{P}(A) = 0$ if $A \in \mathcal{A}$ is countable and $\mathbf{P}(A) = 1$ otherwise.

   (a) Show that $\mathcal{A}$ is a $\sigma$-algebra on $\Omega$.

   (b) Show that $\mathbf{P}$ is a probability measure on $(\Omega, \mathcal{A})$.

   Solution.

   (a) We check the axioms for a $\sigma$-algebra. $\emptyset \in \mathcal{A}$ because it's countable. $[0,1] \in \mathcal{A}$ because its complement is $\emptyset$ which is countable.

   Suppose $A \in \mathcal{A}$. We need to check that $A^c \in \mathcal{A}$. Since $A \in \mathcal{A}$, either $A$ is countable or $A^c$ is countable. If $A$ is countable, then the complement of $A^c$ which is $A$ is countable, so $A^c \in \mathcal{A}$. If $A^c$ is countable, then $A^c \in \mathcal{A}$ by definition.

   Suppose $A_n \in \mathcal{A}$, where $n = 1, 2, \ldots$. We need to check that $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$. Since $A_n \in \mathcal{A}$, either all $A_n$ are countable, or there is $A_{n_0}$ which has a countable complement. If all $A_n$ are countable, then $\cup_{n=1}^{\infty} A_n$ is countable, so $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$. If there exists an $A_{n_0}$ with countable complement, then noting that $(\cup_{n=1}^{\infty} A_n)^c \subset A_{n_0}^c$, it follows that the complement of $\cup_{n=1}^{\infty} A_n$ is countable, so $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$.

   (b) Since $[0,1]$ is not countable, $\mathbf{P}([0,1]) = 1$. Suppose $A_n$ are pairwise disjoint elements of $\mathcal{A}$. We need to check that $\mathbf{P}(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbf{P}(A_n)$. Either all $A_n$ are countable, or there exists $A_{n_0}$ with a countable complement. If all $A_n$ are countable, then $\mathbf{P}(A_n) = 0$, so $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = 0$. Also $\cup_{n=1}^{\infty} A_n$ is countable, so $\mathbf{P}(\cup_{n=1}^{\infty} A_n) = 0$. We conclude that $\mathbf{P}(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbf{P}(A_n)$.

   If $A_{n_0}$ has a countable complement, then by the fact that the $A_n$ are pairwise disjoint, $A_n \subset A_{n_0}^c$ for all $n \neq n_0$. It follows that the $A_n$ for $n \neq n_0$ are countable. Hence, $\mathbf{P}(A_{n_0}) = 1$ and $\mathbf{P}(A_n) = 0$ for $n \neq n_0$, and $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = 1$. In addition, $\cup_{n=1}^{\infty} A_n$ has a countable complement, so $\mathbf{P}(\cup_{n=1}^{\infty} A_n) = 1$ as required.

2. Consider an infinite sequence of independent coin flips such that the probability of a head appearing on the $n$th flip is $p_n$. Denote $A$ the event that only finitely many heads appear.

   (a) Show that $A$ is a tail event. Deduce that either $\mathbf{P}(A) = 0$ or $\mathbf{P}(A) = 1$.

   (b) Find necessary and sufficient conditions on the $p_n$ for $\mathbf{P}(A) = 0$.

   Solution.

   (a) Denote $X_n$ a random variable that equals 1 if the $n$-th flip results in a head and equals 0 if it results in a tail. Let $\mathcal{A}_n$ be the $\sigma$-algebra generated by $X_n, X_{n+1}, \ldots$. The tail $\sigma$-algebra is $\mathcal{A} = \cap_{n=1}^{\infty} \mathcal{A}_n$. We need to show that $A \in \mathcal{A}$. Consider the event $A^c$ which consists in infinitely many heads appearing. We know that $A^c = \limsup_{n\to\infty}\{X_n = 1\} = \cap_{n=1}^{\infty} \cup_{k=n}^{\infty} \{X_k = 1\}$. The event $\cup_{k=n}^{\infty}\{X_k = 1\}$ belongs to $\mathcal{A}_n$ so it belongs to $\mathcal{A}_m$ if $n \geq m$. Therefore, $\cap_{n=m}^{\infty} \cup_{k=n}^{\infty} \{X_k = 1\}$ also belongs to $\mathcal{A}_m$. Since $\cap_{n=1}^{\infty} \cup_{k=n}^{\infty} \{X_k = 1\} = \cap_{n=m}^{\infty} \cup_{k=n}^{\infty} \{X_k = 1\}$, we conclude that $\cap_{n=1}^{\infty}\cup_{k=n}^{\infty}\{X_k = 1\}$ belongs to $\mathcal{A}_m$ for arbitrary $m$. Hence, this event belongs to all $\sigma$-algebras $\mathcal{A}_m$, which means that it belongs to its intersection $\cap_{m=1}^{\infty}\mathcal{A}_m$. Thus, $A^c$ belongs to the tail $\sigma$-algebra. It follows that $A$ belongs to the tail $\sigma$-algebra because of the definition of a $\sigma$-algebra.

Finally, since the random variables $X_n$ are independent, by Kolmogorov's 0-1 law the tail $\sigma$-algebra consists only of events of probability zero or one. Hence, $A$ either is of probability 1 or of probability 0.

(b) The event $A^c$ consists in there appearing infinitely many heads. By the Borel-Cantelli lemma, $\mathbf{P}(A^c) = 0$ if and only if $\sum_{n=1}^{\infty} p_n < \infty$. Thus, $\mathbf{P}(A) = 0$ if and only if $\sum_{n=1}^{\infty} p_n = \infty$.

3. Let $X$ and $Y$ be independent random variables with $XY = 1$ a.s. Show that $X$ and $Y$ are constant a.s.

Solution.

(a) Suppose, in addition, that $X \geq 0$ and $Y \geq 0$ a.s. and that $\mathbf{E}X < \infty$ and $\mathbf{E}Y < \infty$. By independence, $1 = \mathbf{E}(XY) = \mathbf{E}X\mathbf{E}Y$, so $\mathbf{E}X\mathbf{E}Y = 1$. Suppose, $\mathbf{P}(X \neq \mathbf{E}X) > 0$. Then $\mathbf{P}(X > \mathbf{E}X) > 0$.

Indeed, if $\mathbf{P}(X > \mathbf{E}X) = 0$, then

$$\begin{aligned}\mathbf{E}X &= \mathbf{E}(X\mathbf{1}_{\{X=\mathbf{E}X\}}) + \mathbf{E}(X\mathbf{1}_{\{X>\mathbf{E}X\}}) + \mathbf{E}(X\mathbf{1}_{\{X<\mathbf{E}X\}}) \\ &= (\mathbf{E}X)\mathbf{P}(X = \mathbf{E}X) + \mathbf{E}(X\mathbf{1}_{\{X<\mathbf{E}X\}})\end{aligned}$$

Hence, $(\mathbf{E}X)\mathbf{P}(X \neq \mathbf{E}X) = \mathbf{E}(X\mathbf{1}_{\{X<\mathbf{E}X\}})$. The probability on the left is $\mathbf{P}(X < \mathbf{E}X)$ because we supposed that $\mathbf{P}(X > \mathbf{E}X) = 0$. Thus, $(\mathbf{E}X)\mathbf{P}(X < \mathbf{E}X) = \mathbf{E}(X\mathbf{1}_{\{X<\mathbf{E}X\}})$, or $\mathbf{E}((X-\mathbf{E}X)\mathbf{1}_{\{X<\mathbf{E}X\}}) = 0$. The random variable $(X-\mathbf{E}X)\mathbf{1}_{\{X<\mathbf{E}X\}})$ is nonpositive. Since its expectation is zero it equals zero a.s. Hence, $\mathbf{P}(X < \mathbf{E}X) = 0$, which together with $\mathbf{P}(X > \mathbf{E}X) = 0$ contradicts the supposition that $\mathbf{P}(X \neq \mathbf{E}X) > 0$.

Since $XY = \mathbf{E}X\mathbf{E}Y$ a.s., we have that $Y < \mathbf{E}Y$ on the event $\{X > \mathbf{E}X\}$. We have that $\mathbf{P}(Y \neq \mathbf{E}Y) \geq \mathbf{P}(X > \mathbf{E}X) > 0$. By the argument used for proving that $\mathbf{P}(X > \mathbf{E}X) > 0$ we conclude that $\mathbf{P}(Y > \mathbf{E}Y) > 0$. By independence, $\mathbf{P}(X > \mathbf{E}X)\mathbf{P}(Y > \mathbf{E}Y) = \mathbf{P}(X > \mathbf{E}X, Y > \mathbf{E}Y)$. Hence, the latter probabilty is positive. However, on the event $\{X > \mathbf{E}X, Y > \mathbf{E}Y\}$ we have that $XY > \mathbf{E}X\mathbf{E}Y$. Thus, $\mathbf{P}(XY > \mathbf{E}X\mathbf{E}Y) = \mathbf{P}(XY > 1) > 0$. The obtained contradiction completes the proof.

(b) Consider the general case. Let us prove that for arbitrary constant $a > 0$, $\mathbf{P}(|X| \geq a)$ is either zero or one. Suppose $\mathbf{P}(|X| \geq a) > 0$. By the fact that $XY = 1$ a.s. and independence, we have for $\epsilon > 0$

$$\begin{aligned}\mathbf{P}(|X| \geq a)\mathbf{P}(|X| < a) &= \mathbf{P}(|X| \geq a)\mathbf{P}(|Y| > 1/a) \\ &= \mathbf{P}(|X| \geq a, |Y| > 1/a).\end{aligned}$$

The latter event is of probability zero since on it $|XY| > 1$. Thus, if $\mathbf{P}(|X| \geq a) > 0$, then $\mathbf{P}(|X| < a) = 0$. Hence, if $\mathbf{P}(|X| \geq a) > 0$, then $\mathbf{P}(|X| \geq a) = 1 - \mathbf{P}(|X| < a) = 1$.

Let $\hat{a} = \sup\{a : \mathbf{P}(|X| \geq a) = 1\}$. Since $\mathbf{P}(|X| \geq a) \to 0$ as $a \to \infty$, we have that $\hat{a} < \infty$. By definition, for arbitrary $\epsilon > 0$, $\mathbf{P}(|X| \geq \hat{a} + \epsilon) < 1$, so $\mathbf{P}(|X| \geq \hat{a} + \epsilon) = 0$. On letting $\epsilon \to 0$, $\mathbf{P}(|X| > \hat{a}) = 0$. Also $\mathbf{P}(|X| \geq \hat{a} - \epsilon) = 1$, so $\mathbf{P}(|X| \geq \hat{a}) = 1$. It follows that $\mathbf{P}(|X| = \hat{a}) = 1$. In particular, $\hat{a} > 0$. Therefore, $|Y| = 1/|X| = 1/\hat{a}$ a.s.

Next, by independence

$$\mathbf{P}(X = \hat{a})\mathbf{P}(X = -\hat{a})$$
$$= \mathbf{P}(X = \hat{a})\mathbf{P}(Y = -1/\hat{a}) = \mathbf{P}(X = \hat{a}, Y = -1/\hat{a}).$$

The latter event is of probability zero. Hence, on recalling that $\mathbf{P}(X = \hat{a}) + \mathbf{P}(X = -\hat{a}) = 1$, either $X = \hat{a}$ a.s. or $X = -\hat{a}$ a.s. Accordingly, either $Y = 1/\hat{a}$ a.s. or $Y = -1/\hat{a}$ a.s.

4. Let $(X, Y)$ be a bivariate normal random variable with mean $(0, 0)$. Show that $X/Y$ has a Cauchy distribution.

   Solution. Let $g$ be a bounded measurable function. We have bu a change of variables

$$\mathbf{E}g(X/Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x/y) f_{XY}(x, y)\, dxdy$$
$$= \int_{0}^{\infty}\int_{-\infty}^{\infty} g(z) f_{XY}(x, y)\, y\, dzdy + \int_{-\infty}^{0}\int_{\infty}^{-\infty} g(x/y) f_{XY}(yz, y) y\, dzdy$$
$$= 2\int_{0}^{\infty}\int_{-\infty}^{\infty} g(z) f_{XY}(yz, y)\, y\, dzdy = \int_{-\infty}^{\infty} \big(2\int_{0}^{\infty} f_{XY}(yz, y)\, y\, dy\big) g(z)\, dz\,.$$

   Hence,

$$f_{X/Y}(z) = 2\int_{0}^{\infty} f_{XY}(yz, y)\, y\, dy\,.$$

   Let $Q$ be the covariance matrix. Since

$$f_{XY}(x, y) = \frac{1}{2\pi(\det Q)^{1/2}} e^{-(x,y)Q^{-1}(x,y)^T/2},$$

   we obtain

$$f_{X/Y}(z) = \frac{1}{\pi(\det Q)^{1/2}} \int_{0}^{\infty} e^{-y^2(z,1)Q^{-1}(z,1)^T/2}\, y\, dy$$
$$= \frac{1}{2\pi(\det Q)^{1/2}} \int_{0}^{\infty} e^{-u(z,1)Q^{-1}(z,1)^T/2}\, du = \frac{1}{\pi(\det Q)^{1/2}(z,1)Q^{-1}(z,1)^T}\,.$$

If

$$Q = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix},$$

then $\det Q = \sigma_X^2\sigma_Y^2(1-\rho^2)$ and

$$Q^{-1} = \frac{1}{\sigma_X^2\sigma_Y^2(1-\rho^2)}\begin{pmatrix} \sigma_Y^2 & -\rho\sigma_X\sigma_Y \\ -\rho\sigma_X\sigma_Y & \sigma_X^2 \end{pmatrix}.$$

Hence,

$$\begin{aligned} f_{X/Y}(z) &= \frac{\sigma_X\sigma_Y\sqrt{1-\rho^2}}{\pi(z^2\sigma_Y^2 - 2z\rho\sigma_X\sigma_Y + \sigma_X^2)} = \frac{\sigma_X\sigma_Y\sqrt{1-\rho^2}}{\pi((z\sigma_Y - \rho\sigma_X)^2 + \sigma_X^2(1-\rho^2))} \\ &\qquad \frac{\sigma_X\sigma_Y\sqrt{1-\rho^2}}{\pi\sigma_Y^2((z-\rho\sigma_X/\sigma_Y)^2 + \sigma_X^2(1-\rho^2)/\sigma_Y^2)} \\ &= \frac{1}{\pi(\sigma_X/\sigma_Y)\sqrt{1-\rho^2}(1 + (z-\rho\sigma_X/\sigma_Y)^2/(\sigma_X^2(1-\rho^2)/\sigma_Y^2))} \end{aligned}$$

5. Let $(X_1, X_2, \dots, X_n)$ be an $n$-dimensional Gaussian random variable with zero mean. Let $\mathcal{L}$ denote the collection of all linear combinations with deterministic weights of the $X_i$. In the vector space $\mathcal{L}$ define the inner product $X \cdot Y = \mathbf{E}XY$. Apply the Gram–Schmidt procedure in order to orthogonalize the $X_i$ with respect to this inner product. Show that the random variables $Y_1, \dots, Y_n$ thus obtained are Gaussian and independent.

   Solution. According to the Gram-Schmidt process,

$$\begin{aligned} Y_1 &= X_1, \\ Y_2 &= X_2 - \frac{X_2 \cdot Y_1}{Y_1 \cdot Y_1} Y_1, \\ &\dots, \\ Y_k &= X_k - \sum_{i=1}^{k-1} \frac{X_k \cdot Y_i}{Y_i \cdot Y_i} Y_i, \\ &\dots, \\ Y_n &= X_n - \sum_{i=1}^{n-1} \frac{X_k \cdot Y_i}{Y_i \cdot Y_i} Y_i. \end{aligned}$$

   Thus, for each $k$, $Y_k$ is a linear combination of $X_1, \dots, X_k$ with deterministic weights. Hence, the vector $(Y_1, \dots, Y_n)$ is a linear transformation of $(X_1, \dots, X_n)$, so it is Gaussian. An easy induction argument shows that $\mathbf{E}Y_k = 0$ for $k = 1, 2, \dots, n$. Finally, since by Gram-Schmidt $Y_i \cdot Y_j = 0$ for $i \neq j$, it follows that $\mathbf{E}Y_iY_j = 0$. Thus, the covariance matrix of $(Y_1, \dots, Y_n)$ is diagonal, so they are independent random variables.