Full Name: ______________________

Student No: ______________________

Grade:

For each answer, **provide a short argument**.

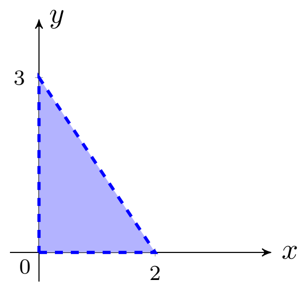
1. (5 points) Let X and Y be two random variables with joint pdf

$$f(x,y) := \begin{cases} k & \text{if } x > 0,\ y > 0 \text{ and } \dfrac{1}{2}x + \dfrac{1}{3}y < 1, \\ 0 & \text{otherwise,} \end{cases}$$

where k is a real number.

(a) Find the value of k.

Solution: From the pdf, one can observe that the random point (X,Y) is uniformly distributed over the triangle with vertices $(0,0)$, $(2,0)$ and $(0,3)$.



Indeed, the joint pdf is constant inside and zero outside this triangle.
The probabilities must add up to 1, hence the integral of f over the entire plane, $\iint_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$, must be equal to 1. However, this integral is nothing but $k \times \langle\text{area of the triangle}\rangle$. Therefore, $k \times \frac{1}{2}(2 \times 3) = 1$, which implies $\boxed{k = 1/3}$.

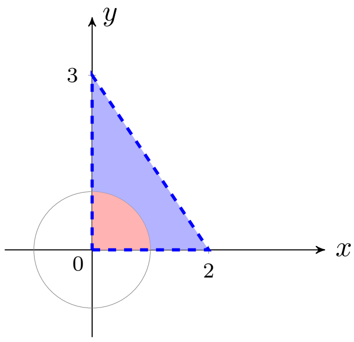
(b) Are X and Y independent?

Solution: The two variables are not independent because knowing the value of one gives us some information on the value of the other. For instance, each of the two events $X > 1$ and $Y > 1.5$ has a positive probability, but if one of them happens, the other cannot. More specifically,

$$\mathbb{P}(X > 1 \text{ and } Y > 1.5) = 0 \neq \frac{1}{4} \times \frac{1}{4} = \mathbb{P}(X > 1)\,\mathbb{P}(Y > 1.5)\ .$$

(c) What is the probability that $X^2 + Y^2 < 1$?

Solution: The event $X^2 + Y^2 < 1$ is the event that the point (X,Y) falls inside the disk of radius 1 centered at the origin.



Denoting the disk $\{(x,y) : x^2+y^2 < 1\}$ by D and the triangle $\{(x,y) : x > 0 \text{ and } y > 0 \text{ and } \frac{1}{2}x + \frac{1}{3}y < 1\}$ by T, we can write

$$\mathbb{P}\big((X,Y) \text{ falls in } D\big) = \frac{\langle \text{area of } D \cap T\rangle}{\langle \text{area of } T\rangle} = \frac{\pi/4}{(2\times 3)/2} = \boxed{\pi/12}\,.$$

Equivalently, to obtain $\mathbb{P}\big((X,Y) \text{ falls in } D\big)$, we can integrate the joint pdf over D, but this precisely gives $k \times \langle \text{are of } D \cap T\rangle = \pi/12$.

(d) Find the expected value of X.

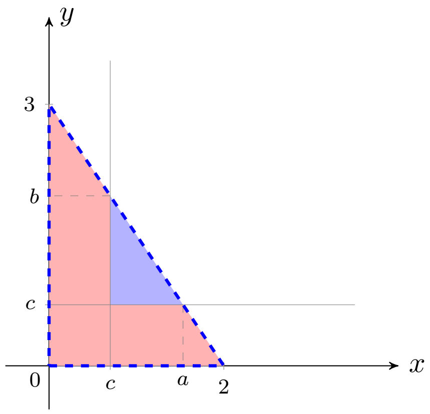
Solution: We have

$$\begin{aligned}
\mathbb{E}[X] &= \iint_{\mathbb{R}^2} f(x,y)\cdot x \; \mathrm{d}x\,\mathrm{d}y \\
&= \int_{x=0}^{2}\int_{y=0}^{3-\frac{3}{2}x} kx \; \mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=0}^{2} \frac{1}{3}x\Big(3-\frac{3}{2}x\Big)\; \mathrm{d}x \\
&= \Big[\frac{1}{2}x^2 - \frac{1}{6}x^3\Big]_{x=0}^{x=2} \\
&= 2-\frac{4}{3} = \boxed{2/3}\,.
\end{aligned}$$

Alternatively, we can first find the marginal distribution of X and use that to calculate $\mathbb{E}[X]$.

(e) Find the cdf of $Z := \min(X,Y)$. [*Hint:* First identify the possible values of Z.]

Solution: First, note that since $X > 0$ and $Y > 0$, we also have $Z > 0$. Moreover, since $1 > \frac{1}{2}X + \frac{1}{3}Y$, we also have $1 > \frac{1}{2}Z + \frac{1}{3}Z = \frac{5}{6}Z$, that is, $Z < 6/5$. Thus, Z can only take values within the interval $(0, 6/5)$.
Let $F_Z(c) := \mathbb{P}(Z \le c)$ be the cdf of Z. Clearly, $F_Z(c) = 0$ for $c \le 0$ and $F_Z(c) = 1$ for $c \ge 6/5$. Let $0 < c < 6/5$. Note that $Z \le c$ if and only if either $X \le c$ or $Y \le c$.



Thus, $\mathbb{P}(Z \leq c) = \langle\text{area of the red region}\rangle / \langle\text{area of the big triangle}\rangle$.
To find the area of the red region, we need a and b. The similarity between a pair of triangles (which?) implies that $a/2 = (3-c)/3$, hence $a = \frac{2}{3}(3-c)$. Likewise, the similarity between another pair of triangles (which?) implies that $b/3 = (2-c)/2$, hence $b = \frac{3}{2}(2-c)$. It follows that

$$\begin{aligned}
\langle\text{area of the red region}\rangle &= \langle\text{area of the big triangle}\rangle - \langle\text{area of the blue triangle}\rangle \\
&= \frac{1}{2}(3 \times 2) - \frac{1}{2}(a-c)(b-c) \\
&= 3 - \frac{1}{2}\left(\frac{2}{3}(3-c) - c\right)\left(\frac{3}{2}(2-c) - c\right) \\
&= \frac{5}{12}c(12-5c) \; .
\end{aligned}$$

Therefore, $\mathbb{P}(Z \leq c) = \frac{5}{36}c(12-5c)$.
In summary,

$$F_Z(c) = \mathbb{P}(Z \leq c) = \boxed{\begin{cases} 0 & \text{if } c < 0, \\ \dfrac{5}{36}c(12-5c) & \text{if } 0 \leq c \leq 6/5, \\ 1 & \text{if } 6/5 < 0. \end{cases}}$$

As a sanity check, we may observe that $\frac{5}{36}c(12-5c) = 0$ when $c = 0$ and $\frac{5}{36}c(12-5c) = 1$ when $c = 6/5$, which is consistent with the fact that Z is a continuous random variable.

2. (3 points) Let $U_1, U_2, \ldots, U_{1000}$ be independent, identically distributed random variables, each with uniform distribution over the interval $[-\pi, \pi]$. Let

$$\begin{aligned}
S &:= \sin(U_1) + \sin(U_2) + \cdots + \sin(U_{1000}) \; , \\
M &:= \frac{1}{1000}\Big[\sin(U_1) + \sin(U_2) + \cdots + \sin(U_{1000})\Big] \; .
\end{aligned}$$

(a) What is the approximate value of M? [*Hint:* Consider $X_i := \sin(U_i)$ for $i = 1, 2, \ldots, 1000$.]

Solution: For $i = 1, 2, \ldots, 1000$, let $X_i := \sin(U_i)$. Note that $X_1, X_2, \ldots, X_{1000}$ are independent, identically distributed random variables. Hence, by the law of large numbers, their average, M, is approximately equal to the expected value of each. For each i, we have

$$\mathbb{E}[X_i] = \mathbb{E}[\sin(U_i)] = \int_{-\pi}^{\pi} \frac{1}{2\pi} \sin(u) \, \mathrm{d}u = 0 \; .$$

Hence, $\boxed{M \approx 0}$.
We can bound the error in this approximation using Chebyshev's inequality.

(b) Find the (exact) value of $\mathbb{P}(S > 0)$.

Solution: We can use symmetry.
Clearly, $\mathbb{P}(S < 0) + \mathbb{P}(S = 0) + \mathbb{P}(S > 0) = 1$.
By symmetry, $\mathbb{P}(S < 0) = \mathbb{P}(S > 0)$. Namely, for each i, the distribution of U_i is symmetric around 0. Since $\sin(-u) = -\sin(u)$, the distribution of $X_i := \sin(U_i)$ is also symmetric around 0. It follows that the distribution of $S = \sum_{i=1}^{1000} X_i$ is also symmetric around 0. In particular, $\mathbb{P}(S > 0) = \mathbb{P}(S < 0)$.
On the other hand, the U_is, the X_is, and S are all continuous random variables, hence $\mathbb{P}(S = 0) = 0$.
It follows that $2\,\mathbb{P}(S > 0) = 1$, hence $\boxed{\mathbb{P}(S > 0) = 1/2}$.

(c) Find the approximate value of $\mathbb{P}(S > 5)$. [*Hint:* Recall the identity $\sin(\theta)^2 = \frac{1}{2}\big[1 - \cos(2\theta)\big]$.]

Solution: Since $X_i := \sin(U_i)$, for $i = 1, 2, \ldots, 1000$, are independent and identically distributed, by the central limit theorem, the distribution of S is approximately normal.
By the result of part (a), we have $\mathbb{E}[X_i] = 0$ for each i, hence

$$\mathbb{E}[S] = \sum_{i=1}^{1000} \mathbb{E}[X_i] = 0 \qquad \text{(by the linearity of expectation)}$$

To find the variance of S, we can write

$$\begin{aligned}
\mathbb{V}\mathrm{ar}[X_i] &= \mathbb{E}[X_i^2] - \cancel{(\mathbb{E}[X_i])^2}^{\,0} \\
&= \mathbb{E}[\sin(U_i)^2] \\
&= \int_{-\pi}^{\pi} \frac{1}{2\pi} \sin(u)^2 \,\mathrm{d}u \\
&= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{2}\big[1 - \cos(2u)\big] \,\mathrm{d}u && \text{(a trigonometric identity)} \\
&= \frac{1}{2\pi} \left[\frac{1}{2}u - \frac{1}{4}\sin(2u)\right]_{u=-\pi}^{\pi} \\
&= \frac{1}{2} \,.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{V}\mathrm{ar}[S] &= \sum_{i=1}^{1000} \mathbb{V}\mathrm{ar}[X_i] && \text{(by the independence of } X_i\text{s)} \\
&= 1000 \times \frac{1}{2} \\
&= 500 \,.
\end{aligned}$$

In summary, S is approximately $\mathsf{N}(0, 500)$. Using a computer software such as R (or the app for normal distribution), we can find that

$$\mathbb{P}(S > 5) \approx 1 - F_{\mathsf{N}(0,500)}(5) \approx \boxed{0.41153} \,.$$

Alternatively, since S is approximately $\mathsf{N}(0, 500)$, we can write $S = \sqrt{500}Z + 0$, where Z is approximately standard normal. Denoting the cdf of the standard normal distribution with Φ, we have

$$\begin{aligned}\mathbb{P}(S > 5) = \mathbb{P}(\sqrt{500}Z + 0 > 5) &= \mathbb{P}(Z > 5/\sqrt{500}) \\ &\approx 1 - \Phi(5/\sqrt{500}) = 1 - 0.58847 = \boxed{0.41153}\,.\end{aligned}$$