Full Name: ______________________________

Student No: ______________________________

Grade:

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

1. (4 points) The function

$$F(x) := \mathbb{P}(X \le x) = \begin{cases} 0 & \text{if } x < 0, \\ 0.2\,x^2 & \text{if } 0 \le x < 1, \\ ax + b & \text{if } 1 \le x < 2, \\ 1 & \text{if } 2 \le x. \end{cases}$$

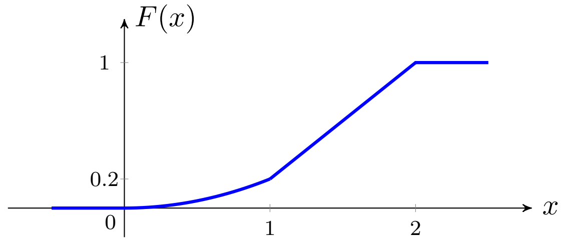
is the cumulative distribution function of a *continuous* random variable X, where a and b are real numbers.

(a) Find the values of a and b, and draw the graph of F.

Solution: The cdf of a continuous RV is a continuous function. The continuity of F at $x = 1$ and $x = 2$ gives

$$0.2 = \lim_{x \uparrow 1} F(x) = F(1) = \lim_{x \downarrow 1} F(x) = a + b\ ,$$
$$2a + b = \lim_{x \uparrow 2} F(x) = F(2) = \lim_{x \downarrow 2} F(x) = 1\ .$$

Solving the two equations $a + b = 0.2$ and $2a + b = 1$, we obtain $\boxed{a = 0.8}$ and $\boxed{b = -0.6}$.



(b) What is the probability that $0.5 \le X \le 2.5$?

Solution: We have

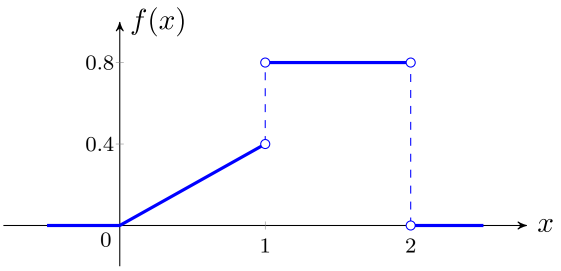
$$\begin{aligned} \mathbb{P}(0.5 \le X \le 2.5) &= \mathbb{P}(X \le 2.5) - \mathbb{P}(X < 0.5) \\ &= \mathbb{P}(X \le 2.5) - \lim_{a \uparrow 0.5} \mathbb{P}(X < a) \\ &= F(2.5) - \lim_{a \uparrow 0.5} F(a) \\ &= 1 - 0.2 \times 0.5^2 \\ &= \boxed{0.95}\ . \end{aligned}$$

(c) Find the probability density function of X.

Solution: The pdf of X is the derivative of its cdf, that is,

$$f(x) = \frac{\mathrm{d}}{\mathrm{d}x}F(x) = \begin{cases} 0 & \text{if } x < 0, \\ 0.4x & \text{if } 0 \le x < 1, \\ 0.8 & \text{if } 1 < x < 2, \\ 0 & \text{if } 2 < x. \end{cases}$$

Note that F is not differentiable at $x = 1$ and $x = 2$, hence f is not defined at those values. It is however differentiable at $x = 0$, despite having different expressions for $x < 0$ and $0 \le x < 1$.



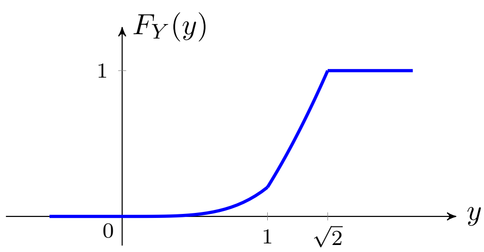
(d) Find the probability density function of $Y := \sqrt{X}$.

Solution: First, note that the possible values of X are $[0, 2]$, hence $\sqrt{X}$ is meaningful. We start by finding the cdf of Y. Clearly, Y can only take non-negative values, hence $F_Y(y) = 0$ for $y < 0$. For $y \ge 0$, we have

$$F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}(\sqrt{X} \le y) = \mathbb{P}(X \le y^2) = \begin{cases} 0.2\,y^4 & \text{if } 0 \le y^2 < 1, \\ 0.8y^2 - 0.6 & \text{if } 1 \le y^2 < 2, \\ 1 & \text{if } 2 \le y^2. \end{cases}$$

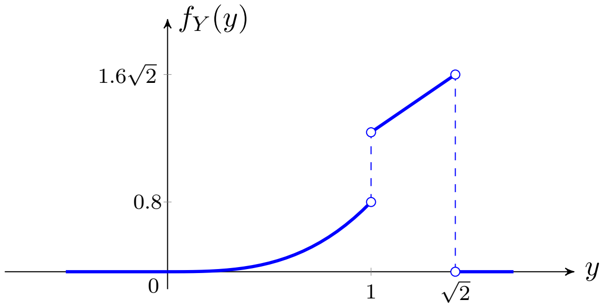
Hence, for arbitrary y, we obtain

$$F_Y(y) = \begin{cases} 0 & \text{if } y < 0, \\ 0.2\,y^4 & \text{if } 0 \le y < 1, \\ 0.8y^2 - 0.6 & \text{if } 1 \le y < \sqrt{2}, \\ 1 & \text{if } \sqrt{2} \le y. \end{cases}$$



Differentiating $F_Y(y)$, we obtain the pdf of Y:

$$f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y}F_Y(y) = \begin{cases} 0 & \text{if } y < 0, \\ 0.8\,y^3 & \text{if } 0 \le y < 1, \\ 1.6y & \text{if } 1 < y < \sqrt{2}, \\ 0 & \text{if } \sqrt{2} < y. \end{cases}$$



2. (1 point) Let M and N be two independent random variables where $M \sim \text{Bin}(n = 4, p = 0.3)$ and $N \sim \text{Bin}(n = 7, p = 0.3)$. Find the distribution of $X := N + M$.
[*Hint:* There is an easy solution that does not require any computation.]

Solution: $X \sim \text{Bin}(n = 11, p = 0.3)$.

The reasoning is similar to the one we used in class to show that the sum of two independent Poisson random variables is again a Poisson random variable, except that we use a Bernoulli process instead of a Poisson process.

Namely, consider a Bernoulli process consisting of a sequence of independent trials each success probability $p = 0.3$. Let M' denote the number of successes among trials $1, 2, 3, 4$, and let N' denote the number of successes among trials $5, 6, 7, 8, 9, 10, 11$. Clearly,

- M' has the same distribution as M,
- N' has the same distribution as N,
- M' and N' are independent, just like M and N.

Therefore, $X' := M' + N'$ has the same distribution as $X = M + N$. But, X' is nothing but the number of successes among the first 11 trials, which has a binomial distribution with parameters $n = 11$ and $p = 0.3$.

An alternative approach is to compute the pmf of X, by noting that X has possible values $0, 1, 2, \ldots, 11$, and for each possible value k, we have

$$p_X(k) = \mathbb{P}(X = k) = \sum_i \mathbb{P}(M = i, N = k - i) \ .$$

This approach is less conceptual, but requires the combinatorial identity $\sum_{i=0}^{n_1} \binom{n_1}{i}\binom{n_2}{k-i} = \binom{n_1+n_2}{k}$.

3. (1 point) Let U, V and W be independent normal random variables, each with parameters $\mu = 3$ and $\sigma^2 = 4$. What is $\mathbb{P}(U + V \leq W)$? It suffices to write your answer in terms of the cdf $\Phi(z)$ of the standard normal distribution.

Solution: The desired probability can be rewritten as

$$\mathbb{P}(U + V \leq W) = \mathbb{P}(\underbrace{U + V - W}_{X} \leq 0) \ .$$

By the stability of the normal distribution, since U, V and W are independent normal random variables, the random variable $X := U + V - W$ is again a normal random variable. Its parameters are

$$\mu_X := \mathbb{E}[X] = \mathbb{E}[U] + \mathbb{E}[V] - \mathbb{E}[W] \qquad \text{(by linearity of expectation)}$$
$$= \mu + \mu - \mu = \mu = 3\,,$$

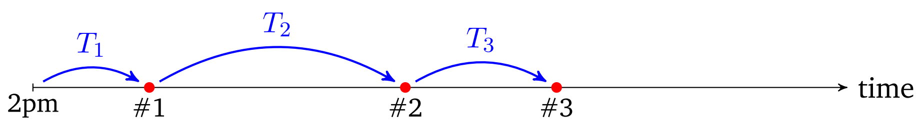
$$\sigma_X^2 := \mathbb{V}\mathrm{ar}[X] = \mathbb{V}\mathrm{ar}[U] + \mathbb{V}\mathrm{ar}[V] + \mathbb{V}\mathrm{ar}[W] \qquad \text{(by independence of } U,\, V \text{ and } W\text{)}$$
$$= \sigma^2 + \sigma^2 + \sigma^2 = 3\sigma^2 = 12\,.$$

Now, recall that every normal random variable is the shifted and scaled version of a standard normal random variable. More specifically, $X = \sigma_X Z + \mu_X$ where Z is a standard normal random variable. Therefore,

$$\mathbb{P}(U + V \leq W) = \mathbb{P}(X \leq 0) = \mathbb{P}(\sigma_X Z + \mu_X \leq 0)$$
$$= \mathbb{P}\left(Z \leq -\mu_X/\sigma_X\right)$$
$$= \Phi\left(-\mu_X/\sigma_X\right) = \Phi(-3/\sqrt{12}) = \boxed{\Phi(-\sqrt{3}/2)} \approx 0.19325\,.$$

4. (1 point) Starting from 2pm when the shop opens, customers arrive at a barber shop according to a Poisson process with a rate of 2 customers per hour. Each customer is served for exactly 30 minutes. If a customer arrives while the shop is empty, they are served immediately. Otherwise, the customer has to wait for their turn. What is the chance that, starting from 2pm, none of the first three customers has to wait before being served?

Solution: Let us call the arrival time of the first customer T_1, the inter-arrival time between the first and the second customer T_2, and the inter-arrival time between the second and the third customer T_3.



Since the customers arrive according to a Poisson process with rate $\lambda = 2$ per hour, the random variables T_1, T_2 and T_3 are independent exponentially distributed random variables, all with the same rate λ.

Clearly, the first customer is served immediately upon arrival. In order for the second and third customers to be served immediately without waiting, it is necessary and sufficient that $T_2 > 0.5$ hour and $T_3 > 0.5$ hour. Therefore, denoting the cdf of Exp$(\lambda = 2)$ by $F_T(t)$, we have

$$\mathbb{P}(\text{none of the first three customers has to wait before being served})$$
$$= \mathbb{P}(T_2 > 0.5 \text{ and } T_3 > 0.5)$$
$$= \mathbb{P}(T_2 > 0.5)\,\mathbb{P}(T_3 > 0.5) \qquad \text{(by independence)}$$
$$= \big(1 - F_T(0.5)\big)\big(1 - F_T(0.5)\big)$$
$$= \mathrm{e}^{-2\times 0.5} \times \mathrm{e}^{-2\times 0.5}$$
$$= \boxed{\mathrm{e}^{-2}}\,.$$