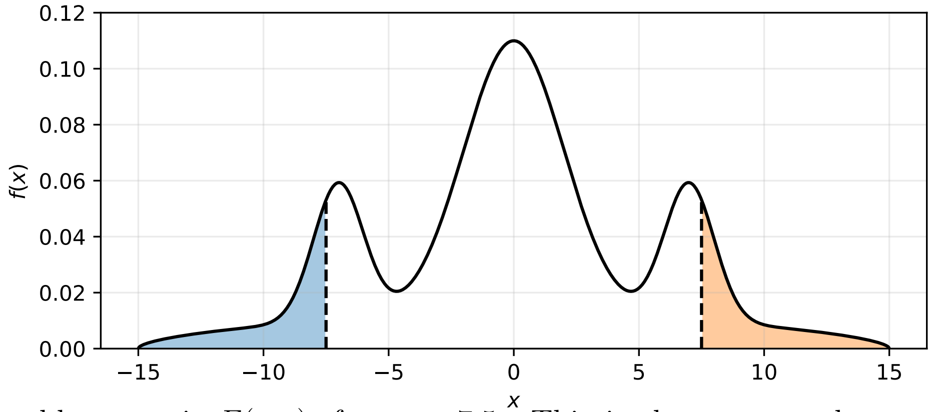
Probability and Statistics for JMC
Solutions 4 — Continuous Random Variables

1. Suppose X is a continuous random variable with density function f which is symmetric around zero, i.e. $\forall x \in \mathbb{R}, f(-x) = f(x)$.
   Show that the cdf satisfies $F(-x) = 1 - F(x)$.

   A picture is worth a thousand words:

   The blue area is $F(-x)$, for $x = 7.5$. This is the same as the orange area because of the symmetry of $f(x)$. The total area under the curve is 1 since it is a (normalized) pdf. The orange area is therefore [the total area under the curve] − [the area under the curve to the left of x] = $1 - F(x)$.
   Or, if you want the above reasoning expressed in symbols:

$$F(-x) = \int_{-\infty}^{-x} f(u)du \quad \underbrace{= \int_{x}^{\infty} f(-s)ds}_{\text{change of variables } s=-u} \quad = \quad \underbrace{\int_{x}^{\infty} f(s)ds}_{\text{symmetry of } f(x)}$$

$$= \int_{-\infty}^{\infty} f(s)ds - \int_{-\infty}^{x} f(s)ds = 1 - F(x).$$

2. Electrons hit a circular plate with radius r_0. Let X be the random variable representing the distance of a particle strike from the center of the plate. Assuming that a particle is equally likely to strike anywhere on the plate,

   (a) for $0 < r < r_0$ find $\mathrm{P}(X < r)$ and write down the full the cumulative distribution function of X, $F_X(r)$ for all r.

   $\mathrm{P}(X < r) = \frac{\text{area of plate within radius } r}{\text{total area of plate}} = \frac{\pi r^2}{\pi r_0^2}$. Therefore,

$$F_X(r) = \begin{cases} 0 & r < 0, \\ (r/r_0)^2 & 0 \le r < r_0, \\ 1 & r \ge r_0. \end{cases}$$

(b) find $\mathrm{P}(r < X < s)$ for $0 \leq r < s \leq r_0$.

$$\mathrm{P}(r < X < s) = F_X(s) - F_X(r) = \tfrac{s^2 - r^2}{r_0^2}.$$

(c) find the probability density function for X, f_X at all r.

$$f_X(r) = \tfrac{dF_X(r)}{dr} = \begin{cases} \frac{2r}{r_0^2} & 0 \leq r \leq r_0, \\ 0 & \text{elsewhere.} \end{cases}$$

(d) calculate the mean distance of a particle strike from the center.

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} r f_X(r) dr = \int_0^{r_0} r \frac{2r}{r_0^2} dr = \frac{2}{3} \frac{r^3}{r_0^2} \bigg|_0^{r_0} = \frac{2}{3} r_0.$$

3. Prove that the mean and variance of an $\mathrm{Exp}(\lambda)$ random variable are $\dfrac{1}{\lambda}$ and $\dfrac{1}{\lambda^2}$ respectively.

The pdf of an $\mathrm{Exp}(\lambda)$ RV is $f(x) = \lambda \exp(-\lambda x)$ for $x > 0$. To do the integrals use integration by parts:

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^{\infty} x \lambda e^{-\lambda x} dx = \int_0^{\infty} x d\left(-e^{-\lambda x}\right) = -x e^{-\lambda x}\big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx.$$

The first term is 0 at both endpoints (use l'Hopital's rule for the $x \to \infty$ case) and the second term is $\int_0^{\infty} \exp(-\lambda x) dx = -(1/\lambda) \exp(-\lambda x)|_0^{\infty} = 1/\lambda$.

For the variance, first calculate $\mathrm{E}(X^2)$ (using integration by parts twice). Then use the fact that for any RV, $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$.

4. Let $X \sim U(0, 1)$. Find the cdf and hence the pdf of the transformed variable $Y = \mathrm{e}^X$.

First, the range of Y is $[1, e]$. For $y \in [1, e]$,

$$F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(e^X \leq y) = \mathrm{P}(X \leq \log y) = \log y,$$

since the cdf of a $U(0, 1)$ RV X is $F_X(x) = x$ for $x \in [0, 1]$. The pdf of Y is therefore,

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{y} \text{ when } 1 \leq y \leq e, \text{ and 0 otherwise.}$$

5. let $X \sim N(\mu, \sigma^2)$, and consider a transformation of variables that defines a new random variable Y, $Y = \dfrac{X - \mu}{\sigma}$. Show that $Y \sim N(0, 1)$.

We will derive a result for a linear transformation of any continuous RV X (see next question for another derivation). Let $Y = aX + b$. It is a one-to-one transformation so we have,

$f_Y(y) dy = f_X(x)|dx|$, where x and dx are related to y and dy by the relation $y = ax + b$. I.e. $x = (y - b)/a$ and $dx = dy/a$. Therefore,

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y - b}{a}\right).$$

Now, for our case $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$ and $Y = \frac{1}{\sigma}X - \frac{\mu}{\sigma}$. Therefore,

$$f_Y(y) = \frac{1}{|1/\sigma|} f_X(\sigma y + \mu) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{y^2}{2}\right],$$

which is the pdf for a $N(0,1)$ RV.

6. Let X be a continuous random variable, with cdf $F_X(x)$ and pdf $f_X(x)$. Let $Y = aX + b$, where $a \neq 0$ and b are constants.

   (a) Considering in turn the two cases $a > 0$ and $a < 0$, use the definition of a cdf to find expressions for the cdf of Y, $F_Y(y)$, in terms of F_X.

   For $a > 0$,

   $$F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(aX + b \leq y) = \mathrm{P}\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right).$$

   For $a < 0$,

   $$F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(aX + b \leq y) = \mathrm{P}\left(X \geq \frac{y-b}{a}\right) = 1 - F_X\left(\frac{y-b}{a}\right).$$

   (b) Using the relationship between a pdf and its cdf, show that the pdf for Y is given by

   $$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right).$$

   For $a > 0$,

   $$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{d}{dy} F_X\left(\frac{y-b}{a}\right) = F_X'\left(\frac{y-b}{a}\right)\frac{1}{a} = \frac{1}{a} f_X\left(\frac{y-b}{a}\right),$$

   using the chain rule and the fact that $F_X'(x) = f_X(x)$. Similarly, for $a < 0$, $f_Y(y) = -\frac{1}{a} f_X\left(\frac{y-b}{a}\right)$. Therefore, for any $a \neq 0$, $f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$.

7. If "area" refers to the area under the curve of the standard normal probability density function ϕ, find the value or values of z such that

   Using the Ch. 6 table of values of Φ, the cdf of a standard normal RV...

   (a) the area between 0 and z is 0.3770.

   There will be two such values, z and $-z$, since the pdf is symmetric around 0. Consider the positive z.
   $\Phi(z) =$ [the area between $-\infty$ and z] $=$ [area between $-\infty$ and 0] $+$ [area between 0 and z. The area between $-\infty$ and 0 is 0.5. Therefore, we want the value of z such that $\Phi(z) = 0.5 + 0.3770 = 0.8770$. The table shows z to be between 1.1 and 1.2. Linearly interpolating between entries in the table we get $z \approx 1.1 + \frac{0.8770-0.864}{0.885-0.864}(1.2-1.1) \approx 1.16$. So the answer is $z = 1.16$ and $z = -1.16$.

   (b) the area to the left of z is 0.8621.

   $\Phi(z) = 0.8621 \Rightarrow z \approx 1.09$

   (c) the area between -1.5 and z is 0.0217.

Area to left of -1.5 is $1 - \Phi(1.5) = 0.067$. So we need to find z such that $\Phi(z) = 0.067 + 0.0217 = 0.0887$. Since this area is less than 0.5, z will be negative. So we find the positive $|z|$ such that the area to the right of $|z|$ is $1 - 0.0887 = 0.9113$. The table tells us it's $|z| = 1.35$. So $z = -1.35$. We also have a solution that is less than -1.5: the solution to $\Phi(-1.5) - \Phi(z) = 0.0217$, which is $z = -1.69$.

8. Find the area under the standard normal pdf
   (a) between $z = 0$ and $z = 1.2$. $\Phi(1.2) - \Phi(0) = 0.3849$
   (b) between $z = -0.68$ and $z = 0$. $\Phi(0) - \Phi(-0.68) = 0.2517$
   (c) between $z = -0.46$ and $z = 2.21$. 0.6637
   (d) between $z = 0.81$ and $z = 1.94$. 0.1828
   (e) to the right of $z = -1.28$. 0.8997

9. You arrive at the bus stop at 10am, knowing that the bus will arrive at some time between 10 and 10:30 with uniform probability.
   (a) What is the probability that you will have to wait longer than 10 minutes?

   Waiting longer than 10 minutes is equivalent to the event "The bus arrives between 10:10 and 10:30," whose probability is (20 mins)/(30 mins) = 2/3.

   (b) If, at 10:15, the bus still has not arrived, what is the probability you will be waiting for at least 10 more minutes?

   The outcomes have been restricted to be between 10:15 and 10:30. There is a uniform probability the bus will arrive at any time in that interval. The probability you are waiting at least 10 more minutes is the probability that the bus arrives between 10:25 and 10:30, which is (5 mins)/(15 mins) = 1/3.

   In probability notation, the original pdf of T, the arrival time of the bus, is

$$f_T(t) = \begin{cases} \frac{1}{30\,\text{mins}} & 10{:}00 \le t \le 10{:}30, \\ 0 & \text{otherwise.} \end{cases}$$

   Now we want $\mathrm{P}(t \ge 10{:}25 \,|\, t \ge 10{:}15)$. Using the definition of conditional probability, this is $\frac{\mathrm{P}(t \ge 10:25 \text{ and } t \ge 10:15)}{\mathrm{P}(t \ge 10:15)} = \frac{\mathrm{P}(t \ge 10:25)}{\mathrm{P}(t \ge 10:15)} = \frac{(5 \text{ mins}/30 \text{ mins})}{(15 \text{ mins } /30 \text{ mins})} = 1/3$.

10. The random variable X has pdf

$$f(x) = \begin{cases} ax + bx^2 & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

   If $\mathrm{E}(X) = 0.6$, find
   (a) $\mathrm{P}\left(X < \frac{1}{2}\right)$

   There are two unknowns, a and b, so we need two equations in order to find them. They are the normalization condition and $\mathrm{E}(X) = 0.6$. From normalization we have,

$$1 = \int_{-\infty}^{\infty} f(x)dx = \int_{0}^{1} (ax + bx^2)dx = \frac{1}{2}a + \frac{1}{3}b,$$

and from expectation we have

$$0.6 = \int\limits_{-\infty}^{\infty} x f(x) dx = \int\limits_{0}^{1} (ax^2 + bx^3) dx = \frac{1}{3}a + \frac{1}{4}b.$$

The solution is $a = 3.6$ and $b = -2.4$.
Then, $\mathrm{P}\left(X < \frac{1}{2}\right) = \int_0^{1/2} f(x)dx = \frac{1}{2}(3.6)(1/2)^2 + \frac{1}{3}(-2.4)(1/2)^3 = 0.35.$

(b) $\mathrm{Var}(X)$

$\mathrm{E}(X^2) = \int\limits_{-\infty}^{\infty} x^2 f(x) dx = \int\limits_{0}^{1} (ax^3 + bx^4) dx = \frac{a}{4} + \frac{b}{5} = 0.42,$
and so $\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2 = 0.42 - 0.6^2 = 0.06.$