1. Let $S = [0, 1]$ equipped with the uniform probability measure (and the Borel $\sigma$-algebra).

(a) Compute $E[X]$, where $X : S \to \mathbb{R}$ is given by $X(s) = \sqrt{s}$.

(b) Compute $E[Y]$, where $Y : S \to \mathbb{R}$ is given by $Y(s) = 0$ if $s \leq \frac{1}{2}$, $Y(S) = 4$ if $s > \frac{1}{2}$.

(c) Compute $E[XY]$, where $X$ and $Y$ are as above. (Yes, I mean the product of the two random variables.) Is $E[XY] = E[X] \cdot E[Y]$?

Solution:

(a) There are two approaches here. One possibility is to come up with the measure that $X$ induces on the real line, find its density, and compute the expectation from there. Let’s compute:

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(s \leq x^2) = \begin{cases} 
0, & x < 0 \\
\sqrt{x}, & 0 \leq x \leq 1 \\
1, & x > 1.
\end{cases}$$

This is differentiable everywhere except at 1, so $X$ has a density:

$$f_X(x) = F_X'(x) = \begin{cases} 
0, & x < 0 \\
2x, & 0 \leq x < 1 \\
0, & x > 1.
\end{cases}$$

The density at 1 is undefined, but that’s ok. We can now compute

$$E[X] = \int_{-\infty}^{\infty} xf_X(x)dx = \int_0^1 x \cdot 2x dx = \frac{2}{3}x^3 |_{x=0}^1 = \frac{2}{3}.$$

The easier solution is to write

$$E[X] = \int_0^1 z^{1/2} \cdot 1 dz = \frac{2}{3}z^{3/2} |_{z=0}^1 = \frac{2}{3}.$$

(b) $Y$ is actually a discrete random variable: $\mathbb{P}(Y = 0) = 1/2$, $\mathbb{P}(Y = 4) = 1/2$. Thus, $E[Y] = 2$.

(c)

$$E[XY] = \int_0^{1/2} z^{1/2} \cdot 0 \cdot 1 dz + \int_{1/2}^{1} z^{1/2} \cdot 4 \cdot 1 dz = \frac{8}{3}z^{3/2} |_{1/2}^1 \approx 1.724.$$

We see that $E[XY] \neq E[X]E[Y]$, and in general we wouldn’t expect expectation to commute with product.

Note that while we can compute the CDF of $XY$, it won’t be continuous (because $XY = 0$ with probability 1/2), thus won’t have a density. So we can’t actually compute $E[XY]$ by integrating with respect to its density.

2. Let $X$ be a normal random variable with mean $\mu$ and variance $\sigma^2$, so that the density of $X$ is

$$f_X(s) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(s-\mu)^2}{\sigma^2}}.$$

Calculate $E[f(X)]$ for the following functions: (Your answer should depend on $\mu$, $\sigma$, $a$, and $b$.)

1
(a) \( f(x) = e^{ax} \), where \( a \) is some real number;

(b) \( f(x) = e^{ax}1_{(b,\infty)}(x) \), where \( a \) and \( b \) are some real numbers. Recall that the indicator function \( 1_{(b,\infty)}(x) = 1 \) if \( x > b \), and 0 otherwise.

(c) \( f(x) = (e^{ax} - b)_+ \), where \( a \) and \( b \) are positive real numbers. Recall that the subscript \((\cdot)_+\) means that we take the positive part of the quantity inside, i.e. \((e^{ax}b)_+ = \max(e^{ax}b, 0)\).

(d) \( f(x) = (b - e^{ax})_+ \), where \( a \) and \( b \) are positive real numbers.

**Solution:**

(a) This is similar to the example we did in class.

\[
E[e^{aX}] = \int_{-\infty}^{\infty} e^{ax} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \, dx
\]

\[
= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma}} \exp \left( -\frac{1}{2} \left( \frac{x^2 - 2\mu x + \mu^2 - 2a\sigma^2 x}{\sigma^2} \right) \right) \, dx
\]

\[
= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma}} \exp \left( -\frac{1}{2} \left( \frac{(x - (\mu + a\sigma^2))^2 + \mu^2 - (\mu + a\sigma^2)^2}{\sigma^2} \right) \right) \, dx
\]

\[
= \exp\left( -\frac{1}{2} \frac{\mu^2 - (\mu + a\sigma^2)^2}{\sigma^2} \right) \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma}} \exp \left( -\frac{1}{2} \left( \frac{(x - (\mu + a\sigma^2))^2}{\sigma^2} \right) \right) \, dx
\]

Notice that the answer depends on \( a\mu \) (the mean of \( aX \)), and \( (a\sigma)^2 \) (the variance of \( aX \)). This is good: the answer should only depend on the moments of \( aX \). (Moments are mean, variance, and higher-order versions of those quantities.) You should be worried if in an expression involving \( aX \) you got terms that looked like \( a^3\mu \) or \( a\sigma^2 \), because this is not the way mean and variance ought to scale with \( a \).

(b) Here and in the rest of the problem, the answer will be less pretty. We can express it in terms of the CDF of a standard normal variable, but we can’t actually evaluate the integral. That said, writing down what the integral should be automatic!

\[
E[e^{aX}1_{(b,\infty)}(X)] = \int_{-\infty}^{\infty} e^{ax}1_{(b,\infty)}(x) \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \, dx
\]

\[
= \int_{-\infty}^{b} 0 \, dx + \int_{b}^{\infty} e^{ax} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \, dx
\]

\[
= \int_{b}^{\infty} e^{ax} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \, dx
\]

This is as far as everyone should have gotten. What I am going to do next is write this integral in terms of \( \Phi \), the CDF of a standard normal \( (N(0,1)) \) random variable. That is:

\[
\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \, dx.
\]

This is a function that most computer software designed for math will compute for you, with arbitrary
precision, and so it is useful to be able to express integrals as above in terms of this standard integral.

\[
\int_{b}^{\infty} e^{ax} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} \, dx
\]

\[
= e^{a\mu + \frac{1}{2} (a\sigma)^2} \int_{b}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} (\frac{x-\mu}{\sigma} + a\sigma)^2} \, dx
\]

Change variables: let \( u = \frac{1}{\sigma} (x - (\mu + a\sigma^2)) = (x - \mu)/\sigma - a\sigma \), then \( du = dx/\sigma \) (so \( dx = \sigma du \)), and we can write

\[
\ldots = e^{a\mu + \frac{1}{2} (a\sigma)^2} \int_{(b-\mu)/\sigma - a\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} u^2} \sigma du
\]

\[
= e^{a\mu + \frac{1}{2} (a\sigma)^2} \int_{(b-\mu)/\sigma - a\sigma}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} u^2} \, du
\]

\[
= e^{a\mu + \frac{1}{2} (a\sigma)^2} \left( 1 - \Phi \left( \frac{b - \mu}{\sigma - a\sigma} \right) \right).
\]

In the last line, we used the fact that \( \Phi \) is the integral of the standard normal pdf from \(-\infty\) up to a point, and the integral from \(-\infty\) to \(+\infty\) must be 1, so the integral from a point to \(+\infty\) is \(1 - \Phi\).

We can do one last simplification, using the fact that the standard normal random variable is symmetric around 0, and therefore \(1 - \Phi(x) = \Phi(-x)\):

\[
e^{a\mu + \frac{1}{2} (a\sigma)^2} \Phi \left( \frac{\mu - b}{\sigma} + a\sigma \right).
\]

The answer again depends on \(a\mu\) and \(a\sigma\), but we also have the \((\mu - b)/\sigma\) term. This comes from the fact that the indicator \(1_{(b,\infty)}(X)\) doesn’t play nicely with rescaling \(X\).

(c)

\[
\mathbb{E}[e^{aX} - b] = \int_{-\infty}^{\infty} \frac{1}{f(x)} \frac{e^{ax} - b}{\sqrt{2\pi}\sigma} \, dx
\]

Let’s look closer at the \((e^{ax} - b)\) term: this is 0 if \(e^{ax} - b \leq 0\), i.e. if \(e^{ax} \leq b\), i.e. if \(x \leq a^{-1} \log b\). (Logarithms are base-e unless specified otherwise.) If \(x > a^{-1} \log b\), then \((e^{ax} - b) = e^{ax} - b\).

Consequently, we rewrite the integral:

\[
\ldots = \int_{-\infty}^{\frac{1}{a} \log b} 0 \, dx + \int_{\frac{1}{a} \log b}^{\infty} \frac{e^{ax} - b}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} \, dx
\]

\[
= \int_{\frac{1}{a} \log b}^{\infty} \frac{e^{ax} - b}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} \, dx
\]

\[
= \int_{\frac{1}{a} \log b}^{\infty} \frac{e^{ax}}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} \, dx
\]

\[
- b \left( \int_{\frac{1}{a} \log b}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\sigma^2}} \, dx \right).
\]

Probability that a \( N(\mu, \sigma) \) random variable will be \( > \frac{1}{a} \log b\)

The nice part is that we know how to compute each of these integrals in terms of \( \Phi \). For the first one, use the previous part of the question, with \( \frac{1}{a} \log b \) instead of \(b\):

\[
e^{a\mu + \frac{1}{2} (a\sigma)^2} \Phi \left( \frac{\mu - \frac{1}{a} \log b}{\sigma} + a\sigma \right).
\]
For the second integral, use 0 instead of $a$ and $\frac{1}{a}\log b$ instead of $b$. Using $a = 0$ means that $e^{ax} = 1$ doesn’t do anything, and we find that the second integral is

$$b\Phi\left(\frac{\mu - \frac{1}{a}\log b}{\sigma}\right).$$

(Careful that the $a$ stays in the expression $\frac{1}{a}\log b$) Thus, we get

$$e^{a\mu + \frac{1}{a}(a\sigma)^2} \Phi\left(\frac{a\mu - \log b}{a\sigma} + a\sigma\right) - b\Phi\left(\frac{a\mu - \log b}{a\sigma}\right)$$

A final touch is to write things in terms of $a\mu$ and $a\sigma$ as much as possible:

$$e^{a\mu + \frac{1}{a}(a\sigma)^2} \Phi\left(\frac{a\mu - \log b}{a\sigma} + a\sigma\right) - b\Phi\left(\frac{a\mu - \log b}{a\sigma}\right).$$

(d) The easiest way to get the answer here is to notice that $(-x)_+ = -x + (x)_+$. (Check this! If $x$ is positive, then $(-x)_+ = 0$ and $x_+ = x$ so it works; if $x$ is negative, then $(-x)_+ = -x$ and $(x)_+ = 0$ so it also works.) Therefore,

$$(b - e^{ax})_+ = b - e^{ax} + (e^{ax} - b)_+.$$ Consequently,

$$E[(b - e^{aX})_+] = b - E[e^{aX}] + E[(e^{ax} - b)_+].$$

We already know how to compute all parts of this, so the answer is

$$b - e^{a\mu + \frac{1}{a}(a\sigma)^2} + e^{a\mu + \frac{1}{a}(a\sigma)^2} \Phi\left(\frac{a\mu - \log b}{a\sigma} + a\sigma\right) - b\Phi\left(\frac{a\mu - \log b}{a\sigma}\right)$$

$$= b(1 - \Phi\left(\frac{a\mu - \log b}{a\sigma}\right)) - e^{a\mu + \frac{1}{a}(a\sigma)^2}(1 - \Phi\left(\frac{a\mu - \log b}{a\sigma} + a\sigma\right))$$

$$= b\Phi\left(\frac{\log b - a\mu}{a\sigma}\right) - e^{a\mu - \frac{1}{a}(a\sigma)^2} \Phi\left(\frac{\log b - a\mu}{a\sigma} - a\sigma\right).$$

Here, we again used the fact that $1 - \Phi(x) = \Phi(-x)$.

3. For each pair of measures $\mu$, $\nu$ below, decide whether $\mu \ll \nu$, $\nu \ll \mu$, both (i.e. $\mu \sim \nu$), or neither. Bonus question: compute the Radon-Nikodym derivatives $d\mu/d\nu$ and $d\mu/d\nu$ (only those that are defined, of course).

(a) Let $S = \{U, D\}$, with the $\sigma$-algebra $\mathcal{F} = 2^S$. Let $\mu$ and $\nu$ be two probability measures, with $\mu(\{U\}) = 0.5$ and $\nu(\{U\}) = 0.6$.

(b) Let $S = \{U, D\}^4$, with the $\sigma$-algebra $\mathcal{F} = 2^S$. Let $\mu$ be a counting measure on $S$, i.e. $\mu$ assigns measure 1 to each one-element set. Let $\nu$ be the uniform probability measure on $S$, i.e. $\nu$ assigns equal measure to each one-element set, and $\nu(S) = 1$.

(c) Let $S = \mathbb{N}$, with the $\sigma$-algebra $\mathcal{F} = 2^S$. Let $\mu(\{n\}) = 1/n; \nu(\{n\}) = 1$ if $n$ is even, and 0 if $n$ is odd.

(d) Let $S = \mathbb{N}$, with the $\sigma$-algebra $\mathcal{F} = \{\emptyset, \{\text{even integers}\}, \{\text{odd integers}\}, \mathbb{N}\}$. Let $\mu$ and $\nu$ be two probability measures, with $\mu(\{\text{even integers}\}) = 0.5$ and $\nu(\{\text{even integers}\}) = 0.6$.

(e) Let $S = \{1, \ldots, 6\}$, with the $\sigma$-algebra $\mathcal{F}$ generated by $\{\text{odd numbers}\} = \{1, 3, 5\}$ and $\{\text{prime numbers}\} = \{2, 3, 5\}$. (What are all the elements of $\mathcal{F}$?) Let $\mu$ be the counting measure (defined on $\mathcal{F}$), i.e. $\mu(A)$ is the number of elements in $A$, and let $\nu$ be the probability measure defined by $\nu(\{3, 5\}) = 1$. 4
(f) Let $S = \mathbb{R}$, with the Borel $\sigma$-algebra. Let $\mu$ be the Lebesgue measure. Let $\nu$ be the measure given by $\nu(A) = \{\text{number of integers in } A\}$. 

(g) Let $S = \mathbb{R}$, with the Borel $\sigma$-algebra. Let $\mu$ be the Lebesgue measure. Let $\nu$ be the law of the random variable $X : [0, 1] \to \mathbb{R}$, the underlying probability measure on $[0, 1]$ is uniform, and $X(s) = 9s - 2$. 

(h) Let $S = \mathbb{R}$, with the Borel $\sigma$-algebra. Let $\mu$ be the Lebesgue measure. Let $\nu$ be the law of the standard normal random variable with mean 0 and variance 1. 

(i) Let $S = \mathbb{R}$, with the Borel $\sigma$-algebra. Let $\mu$ be the law of $X$, and let $\nu$ be the law of $Y$, where $X$ and $Y$ are as follows. Let $\Omega = \{U, D\}^2$ with the uniform probability measure; let $X : \Omega \to \mathbb{R}$ be given by $X(\omega) = \text{number of } U\text{'s in } \omega$, and let $Y : \Omega \to \mathbb{R}$ be given by $Y(\omega) = \text{number of } D\text{'s in } \omega$. (Bonus question: are these random variables measurable with respect to each other?) 

(j) Let $S = \mathbb{R}$, with the Borel $\sigma$-algebra. Let $\mu$ be the law of $X$, and let $\nu$ be the law of $Y$, where $X$ and $Y$ are as follows. Let $\Omega = \{U, D\}^2$ with the uniform probability measure; let $X : \Omega \to \mathbb{R}$ be given by $X(\omega) = \text{number of } U\text{'s in } \omega$, and let $Y : \Omega \to \mathbb{R}$ be given by $Y(\omega) = 3 + \text{number of } D\text{'s in } \omega$. (Bonus question: are these random variables measurable with respect to each other?) 

(k) Let $S = \mathbb{R}$, with the Borel $\sigma$-algebra. Let $\mu$ be the law of $X$, and let $\nu$ be the law of $Y$, where $X$ and $Y$ are as follows. Let $\Omega = \{U, D\}^2$ with the uniform probability measure; let $X : \Omega \to \mathbb{R}$ be given by $X(\omega) = \text{number of } U\text{'s in } \omega$, and let $Y : \Omega \to \mathbb{R}$ be given by $Y(\omega) = 1$. (Bonus question: are these random variables measurable with respect to each other?) 

**Solution:**

(a) The only set with $\mu(A) = 0$ is $A = \emptyset$; this is also the only set with $\nu(A) = 0$. Because the sets of measure 0 coincide for the two measures, the measures are equivalent: $\mu \sim \nu$ (that is, $\mu \ll \nu$ and $\nu \ll \mu$). The Radon-Nikodym derivative is $h(U) = d\mu/d\nu(U) = \mu(U)/\nu(U) = 5/6$, $h(D) = d\mu/d\nu(U) = \mu(D)/\nu(D) = 5/4$. Going the other way, $d\nu/d\mu(U) = 6/5$ and $d\nu/d\mu(D) = 4/5$ (the inverse of the fractions we had before).

(b) Again, the only set with $\mu(A) = 0$ is $A = \emptyset$ (since $A$ must have 0 elements), and the only set with $\nu(A) = 0$ is also $A = \emptyset$. Because the sets of measure zero coincide for the two measures, $\mu \sim \nu$. The Radon-Nikodym derivative is $h(s) = d\mu/d\nu(s) = \mu(s)/\nu(s) = 16$ for every $s \in S$; going the other way, $d\nu/d\mu(s) = \frac{1}{16}$ for every $s \in S$.

(c) The only set with $\mu(A) = 0$ is $A = \emptyset$, but that’s not true for $\nu$. We do have $\nu(\emptyset) = 0$, so $\nu \ll \mu$; but we also have, for example, $\nu(\{1\}) = 0$ while $\mu(\{1\}) > 0$, so $\mu \nsubseteq \nu$. (The symbol $\nsubseteq$ means “is not absolutely continuous with respect to”; it’s a crossed-out $\ll$.) The only Radon-Nikodym derivative that’s defined is $d\nu/d\mu$, and we have $h(n) = d\nu/d\mu(n) = \nu(n)/\mu(n) = n$ if $n$ is even, and 0 if $n$ is odd.

(d) Notice that there are only four measurable sets here, and equivalence of measures has to do with measurable sets. Therefore, the two measures are equivalent, because from the point of view of measurable sets this looks exactly like part (a).

Now, when we go to define $h(n) = d\mu/d\nu(n)$, this is really a function that’s defined on all of $\mathbb{N}$, but it must be constant on each of the two smallest measurable sets. (This is what it means for it to be measurable!) Thus, we have $h(n) = d\mu/d\nu(n) = 5/6$ if $n$ is even, and $= 5/4$ if $n$ is odd. Going the other way, the function $d\nu/d\mu(n) = 6/5$ if $n$ is even, and $= 4/5$ if $n$ is odd.

(e) In this example, the smallest measurable sets are $K = \{1\} = \{1, 3, 5\} \cap \{2, 3, 5\}^c$, $L = \{2\} = \{2, 3, 5\} \cap \{1, 3, 5\}^c$, $M = \{3, 5\} = \{1, 3, 5\} \cap \{2, 3, 5\}$ (3 and 5 will be indistinguishable), and $N = \{4, 6\} = \{1, 3, 5\}^c \cap \{2, 3, 5\}^c$ (4 and 6 will also be indistinguishable). $\mathcal{F}$ thus has 16 elements, which are unions of some of $K$, $L$, $M$, and $N$. (For example, the set $\{3, 4, 5, 6\} = M \cup N$ will be in $\mathcal{F}$.)
The counting measure \( \mu \) still only has one set of \( \mu \)-measure 0: \( \mu(A) = 0 \iff A = \emptyset \). Since \( \nu(\emptyset) = 0 \), we have \( \nu \ll \mu \). On the other hand, there are other sets with \( \nu \)-measure zero: for example, \( \nu(\{4,6\}) = 0 \) (because \( \nu \) is a probability measure, and all of its mass is used up on \( \{3,5\} \)). Thus, \( \mu \not\ll \nu \).

The one Radon-Nikodym derivative that’s defined is \( d\nu/d\mu \), and it must be a function on \( \{1, \ldots, 6\} \) that takes only one value on 3 and 5, and only one value on 4 and 6. (This is what it means to be measurable!) The function in question is

\[
\begin{align*}
h(1) &= \frac{d\nu}{d\mu}(1) = \frac{\nu(\{1\})}{\mu(\{1\})} = 0/1 = 0, \\
h(2) &= \frac{d\nu}{d\mu}(2) = \frac{\nu(\{2\})}{\mu(\{2\})} = 0/1 = 0, \\
h(3) &= \frac{d\nu}{d\mu}(3) = \frac{\nu(\{3,5\})}{\mu(\{3,5\})} = 1/2, \\
h(4) &= \frac{d\nu}{d\mu}(4) = \frac{\nu(\{4,6\})}{\mu(\{4,6\})} = 0/2 = 0, \\
h(5) &= \frac{d\nu}{d\mu}(5) = \frac{\nu(\{3,5\})}{\mu(\{3,5\})} = 1/2, \\
h(6) &= \frac{d\nu}{d\mu}(6) = \frac{\nu(\{4,6\})}{\mu(\{4,6\})} = 0/2 = 0.
\end{align*}
\]

In the discrete situation, we can simply use the smallest measurable set containing the element to compute \( d\nu/d\mu \), such as what we did for \( h(3) \).

(f) In the examples with \( \mathbb{R} \), we don’t know how to characterize all sets of Lebesgue measure 0. (It’s genuinely hard.) But, we don’t need to know all of them. We know \( \mu(\{1\}) = 0 \) while \( \nu(\{1\}) = 1 \), so \( \nu \not\ll \mu \). We also know \( \nu(\{0.1,0.9\}) = 0 \) while \( \mu(\{0.1,0.9\}) = 0.8 \), so \( \mu \not\ll \nu \).

In fact, these measures are as un-equivalent as they could be: we can write \( \mathbb{R} = S_1 \cup S_2 \) where \( S_1 \) and \( S_2 \) are disjoint, \( \mu(S_2) = 0 \) and \( \nu(S_1) = 0 \). That is, the measures life on separate parts of \( \mathbb{R} \). Specifically, take \( S_2 = \mathbb{Z} \), then \( \mu(S_2) = 0 \) because \( S_2 \) is countable; and take \( S_1 = \mathbb{R} - \mathbb{Z} \). Such measures are sometimes called orthogonal.

(g) Here it’s more of an issue that we can’t characterize sets of \( \mu \)-measure 0. Instead, we go the CDF approach: the CDF of \( \nu \) is a continuous function that is 0 to the left of \(-2\), has slope 1/9 on the interval \([-2,7] \), and is 1 to the right of 7. If we could draw the CDF of \( \mu \), we could compare their points of continuity and intervals where they’re flat, but the CDF of \( \mu \) is actually infinite. (\( \mu(\{x\}) = \infty \) for all \( x \).

Nonetheless, because the CDF of \( \nu \) is continuous and differentiable almost everywhere, \( \nu \) has a pdf, which is the density with respect to the Lebesgue measure. In particular, yes, \( \nu \ll \mu \). (There is some handwaving involved here. I will not prove rigorously that if \( \mu(A) = 0 \) then \( \nu(A) = 0 \) in this course, because that would require me to go deeper into the theory of measure and integration than we’ve gone already. The general idea is that \( d\mu = dx \), and because the density of \( \nu \) is bounded, \( \nu(A) = \int_A f_{\nu}(x)dx \) cannot be much larger than \( \mu(A) = \int_A 1dx \).)

For the purposes of testing absolute continuity, you can think of the CDF of the Lebesgue measure as differentiable everywhere with slope 1. Thus, it’s never flat, and never has jumps.

Coming back to the problem at hand: \( \nu \ll \mu \) because the CDF of \( \nu \) is continuous and differentiable almost everywhere. The derivative \( d\nu/d\mu \) is also the pdf of \( X \): it’s 0 outside the interval \([-2,7] \), and \( \frac{1}{9} \) inside the interval. It’s not defined on \( \{2,7\} \), but that’s a set of measure 0.

On the other hand, \( \mu \not\ll \nu \), because \( \nu([8,9]) = 0 \) while \( \mu([8,9]) = 1 \).
(h) Because the standard normal random variable has a pdf (we keep writing it down), \( \nu \ll \mu \). The derivative \( d\nu/d\mu \) is the pdf of the standard normal random variable, i.e.

\[
h(x) = \frac{d\nu}{d\mu}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.
\]

Moreover, because this density is invertible everywhere, we actually have \( \mu \ll \nu \) as well, i.e. \( \mu \sim \nu \). The derivative is

\[
h(x) = \frac{d\mu}{d\nu}(x) = \left( \frac{d\nu}{d\mu}(x) \right)^{-1} = \frac{1}{\sqrt{2\pi}} e^{x^2/2}.
\]

(i) Here we can test equivalence of measures by drawing the CDF. Notice that \( X \) takes on the values 0, 1, 2 with probability 1/4, 1/2, 1/4 each, and so does \( Y \); so the CDFs of \( \mu \) and \( \nu \) are in fact the same, and definitely equivalent to each other. The derivative can be taken to be \( d\mu/d\nu(x) = 1 \) for all \( x \in \mathbb{R} \). Notice that we actually only know that it must be 1 at the points 0, 1, 2 – since \( \mu(\{0, 1, 2\}^c) = 0 = \nu(\{0, 1, 2\}^c) \), we could put absolutely anything for the value of the derivative outside of the set \( \{0, 1, 2\} \).

To see whether \( X \) is \( \sigma(Y) \)-measurable, we need to check whether knowing the value of \( X \) we could tell the value of \( Y \), and vice versa. Indeed, knowing the number of \( U \)'s in a two-letter word tells you the number of \( D \)'s in it, and conversely, so \( X \) is \( \sigma(Y) \)-measurable, and \( Y \) is \( \sigma(X) \)-measurable.

(j) Here, \( X \) takes the values 0, 1, 2 and \( Y \) takes the values 3, 4, 5. Consequently, \( \mu(\{3\}) = 0 \) while \( \nu(\{3\}) > 0 \), so \( \nu \not\ll \mu \); and \( \nu(\{0\}) = 0 \) while \( \mu(\{0\}) > 0 \), so \( \mu \not\ll \nu \).

On the other hand, knowing the number of \( U \)'s still tells me the number of \( D \)'s in the word, so I can compute \( 3 + \) number of \( D \)'s; and conversely, knowing \( 3 + \) number of \( D \)'s, I can compute the number of \( U \)'s. Consequently, \( X \) is \( \sigma(Y) \)-measurable and \( Y \) is \( \sigma(X) \)-measurable.

We see that whether \( X \) is \( \sigma(Y) \)-measurable has nothing to do with whether the laws of \( X \) and \( Y \) are absolutely continuous with respect to each other.

(k) Here, \( X \) takes the values 0, 1, 2 with non-zero probability for each, and \( Y \) only takes on the value 1. This means that \( \nu \ll \mu \) (because the sets of \( \mu \)-measure zero are sets that don’t include any of 0, 1, or 2, and all of these will also have \( \nu \)-measure 0); but \( \mu \not\ll \nu \) (because \( \nu(\{0\}) = 0 \) while \( \mu(\{0\}) > 0 \)).

The one Radon-Nikodym derivative that’s defined is \( d\nu/d\mu \), and we have

\[
h(x) = \frac{d\nu}{d\mu}(x) = \begin{cases} 0, & x = 0 \\ \frac{1}{2} = 2, & x = 1 \\ 0, & x = 2 \\ (don’t know and don’t care), & x \not\in \{0, 1, 2\} \end{cases}.
\]

On the numbers 0, 1, 2 we can treat these variables as discrete, and compute the derivative as \( \nu(\{x\})/\mu(\{x\}) \). Outside that range, we’re looking at \( (\text{slope of CDF of } \nu) / (\text{slope of CDF of } \mu) = 0/0 \), which isn’t defined – but we don’t care, because the complement of \( \{0, 1, 2\} \) has measure 0.