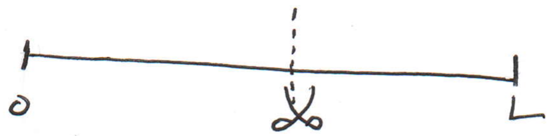
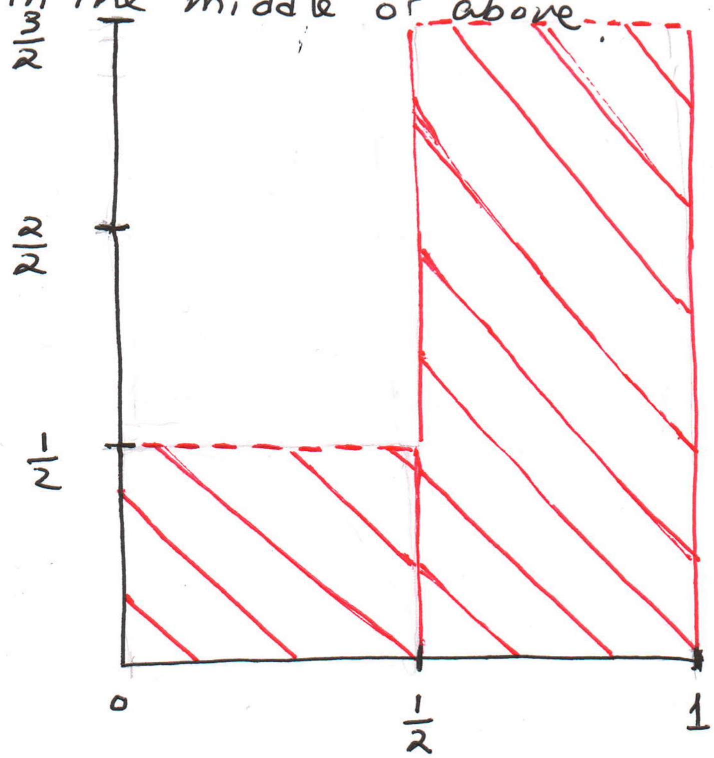
Continuous Random Variables Lecture 1

To see how continuous random variables and their important features might arise, imagine that you are watching a barber cut hair. You have made detailed video recordings of 1000 individual hair strands being cut. Suppose that each of the hair strands are of the same length and are treated by the barber independently.

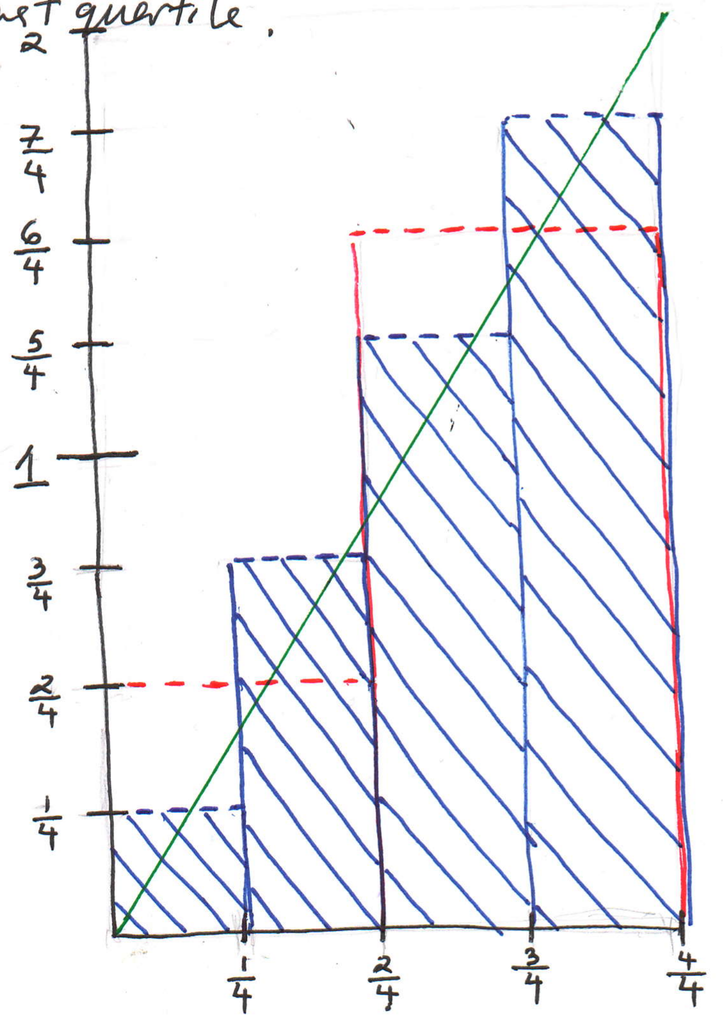


1) of the hair strands observed, $\frac{1}{4}$ appear to have been cut in the middle or below and $\frac{3}{4}$ appear to have been cut in the middle or above.



The height of each red rectangle is adjusted to make its area equal to the corresponding probability.

2) Next it is observed that approximately $\frac{1}{16}$ of the hairs have been cut between the root and $\frac{1}{4}$ length, $\frac{3}{16}$ have been cut between $\frac{1}{4}$ and $\frac{2}{4}$ from the scalp, $\frac{5}{16}$ of the hairs were cut at a point $\frac{2}{4}$ to $\frac{3}{4}$ from the scalp, and $\frac{7}{16}$ of the hairs were cut within the last quartile.



Area of each bar of the histogram equals to corresponding probability.

As we continue to make finer measurements, we generate histograms with a multitude of thin bars. The thickness of each bar is adjusted to match the probability of the cut taking place at a point over the base of the bar.

We might eventually notice that the tips of the bars align along the curve $y = 2x$ (in green).

If X is the exact place where the hair (of normalized length) is cut, then by construction, $P(X \in (a,b)) =$

$= P(a < X < b) = \int_a^b 2x\,dx = b^2 - a^2$. This just means that we are adding the areas of thin bars of a very fine histogram that lie between point a and point b.

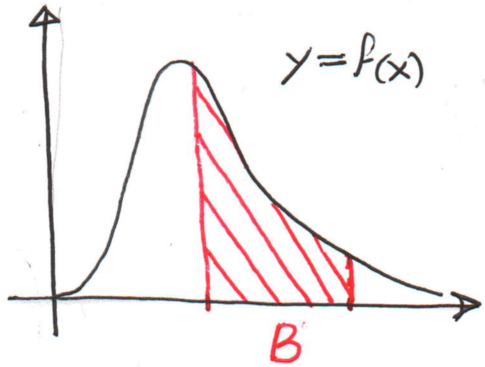
For instance, the probability this barber cuts some particular hair to a size between $\frac{1}{4}$ to $\frac{1}{2}$ of its original length is $P(\frac{1}{4} \leq X \leq \frac{1}{2}) = \int_{\frac{1}{4}}^{\frac{1}{2}} 2x\,dx = (\frac{1}{2})^2 - (\frac{1}{4})^2$

$= \frac{1}{4} - \frac{1}{16} = \frac{1}{4}(1 - \frac{1}{4}) = \frac{3}{16}$.

Since we could conceivably replicate this process of thin histogram construction for other phenomena, we get the following definition.

<u>Def:</u> $X \in (-\infty, \infty)$ is said to be a continuous random variable if there exists a continuous density function $f: \mathbb{R} \longrightarrow [0, \infty)$ such that

$$P(X \in B) = \int_B f(x)\,dx$$ for any (measurable) subset $B \subset \mathbb{R}$.

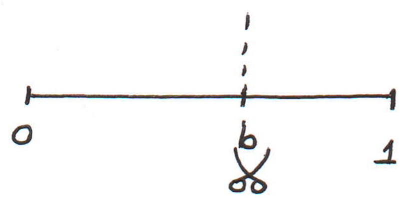


<u>Ex.</u> (The haircut Paradox) A barber whose density function is $f(x) = 2x$; $0 \leq x \leq 1$ comes to cut your hair.

(a) what is the probability the cut is made half-way between the scalp and the tip?

(b) what is the probability the cut is made $\frac{7}{8}$ ths hair lengths from scalp?

(c) what is the probability the cut is made through the point $0 \leq b \leq 1$?



<u>Solution:</u>

(a) $P(X = \frac{1}{2}) = \int_{\frac{1}{2}}^{\frac{1}{2}} 2x\,dx = 0$

(b) $P(X = \frac{7}{8}) = \int_{\frac{7}{8}}^{\frac{7}{8}} 2x\,dx = 0$

(c) In general for any $b \in [0,1]$, $P(X=b) = \int_b^b 2x\,dx$
$= b^2 - b^2 = 0$.

So... the hair will not be cut?

Notice that for a generic random variable Y with density function $f(y)$, then $P(Y=b) = \int_b^b f(y)\,dy = 0$

and so the probability that Y takes any particular value b is equal to 0!

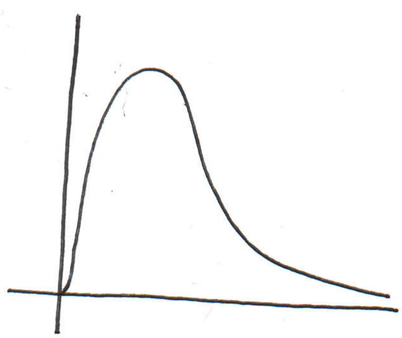
Does this mean that no value will be generated?
We can understand what's happening by recalling that $P(Y=b) = \lim_{n \to \infty} \frac{n(b)}{n}$ where we count the number of instances where b occurs $n(b)$ and divide it by the total number of trials n.
The outcome b can occur infrequently enough to make the limit converge to 0.
Remember $P(Y=b)=0$ does not mean that b never happens!
Another thing to note is that any real number is an infinite sequence of decimals. You cannot physically display all the digits. If only 8 digits are displayed

For example, the probability of getting $X = 0.5000000\ldots$ is $\int_{\frac{1}{2}}^{\frac{1}{2}+10^{-8}} 2x\,dx \approx 2 \cdot \frac{1}{2} \cdot 10^{-8} = 10^{-8} > 0$.

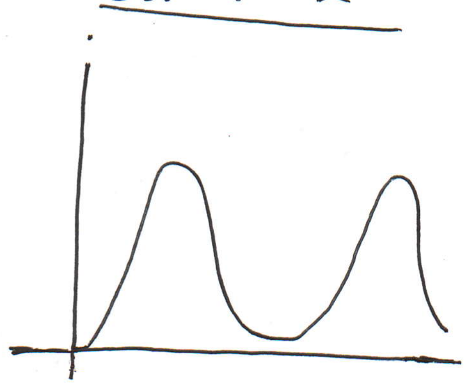
Thus you may also think that $P(X=b) = 0$ means that you will not see all the digits of b.

<u>Ex.</u> A girl wants to get a haircut. She can go to one of 3 barbers. The barbers cut each individual hair strand in accordance with the following density functions.

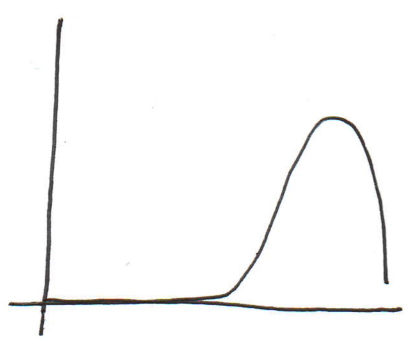
<u>Barber 1</u>



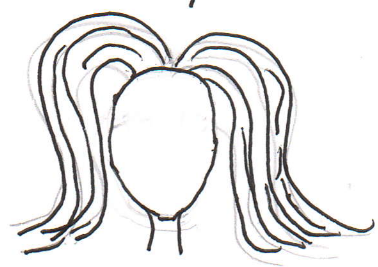
<u>Barber 2</u>



<u>Barber 3</u>



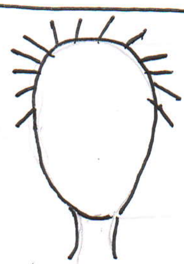
If she enters the barbershop looking like this



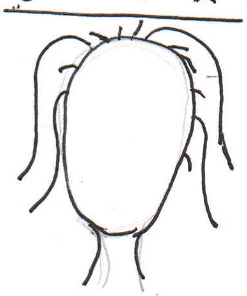
what will she look like any one of the available haircuts?

Solution:

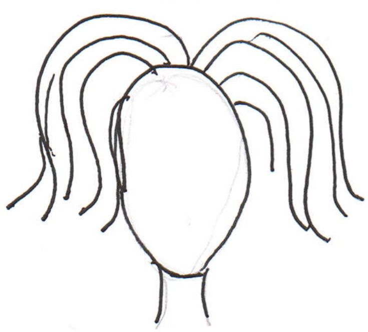
Barber 1



Barber 2



Barber 3



Ex. Suppose that the probability density function (pdf) of X is given by $f(x) = \begin{cases} cx^2 & 0 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$

(a) What must be the value of c ?

(b) Compute $P(1 \leq X \leq 2)$

Solution:

(a) $$1 = \int_{-\infty}^{\infty} f(x)\,dx = \int_0^2 c x^2\,dx = \frac{c}{3} x^3 \Big|_0^2 = c \frac{8}{3}$$

Hence $c = \frac{3}{8}$

(b) $$P(1 \leq X \leq 2) = \int_1^2 f(x)\,dx = \int_1^2 \frac{3}{8} x^2\,dx$$

$$= \frac{1}{8} x^3 \Big|_1^2 = \frac{2^3 - 1^3}{8} = \frac{7}{8}$$

Ex. The amount of time a computer functions before freezing up is a random variable with pdf

$$f(x) = \begin{cases} \lambda e^{-\frac{x}{100}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Find the probability that this computer functions between 50 and 150 hrs.

Solution: $$\int_0^{\infty} \lambda e^{-\frac{x}{100}}\,dx = -\lambda 100 e^{-\frac{x}{100}} \Big|_0^{\infty} = \lambda \cdot 100 = 1$$

So $\lambda = \frac{1}{100}$

$$P(50 \leq X \leq 150) = \int_{50}^{150} \frac{1}{100} e^{-\frac{x}{100}}\,dx = -e^{-\frac{x}{100}} \Big|_{50}^{150} = e^{-\frac{1}{2}} - e^{-\frac{3}{2}}$$

Cumulative distribution of a continuous random variable

If X is a continuous random variable with density f then the cumulative distribution of X (cdf) is

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)\,dx.$$

Recall that by the fundamental theorem of calculus

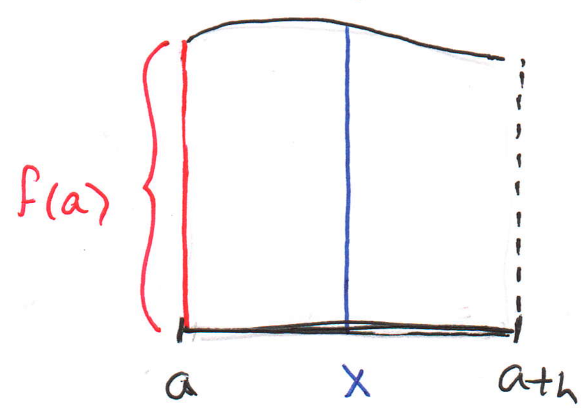
$$f(a) = \frac{d}{da}\int_{-\infty}^{a} f(x)\,dx = \frac{d}{da}F(a)$$

To see this simply observe that $\frac{d}{da}F(a) =$

$$= \lim_{h\to 0}\frac{1}{h}\Big(F(a+h) - F(a)\Big) = \lim_{h\to 0}\frac{1}{h}\Big(\int_{-\infty}^{a+h} f(x)\,dx - \int_{-\infty}^{a} f(x)\,dx\Big) = \lim_{h\to 0}\frac{1}{h}\int_{a}^{a+h} f(x)\,dx = \lim_{h\to 0}\frac{1}{h}f(a)h$$

$$= f(a)$$

where we used the continuity of $f(x)$ as follows:



For $h>0$ small enough and any $x \in [a, a+h]$, $f(x)$ is hardly distinguishable from $f(a)$ and therefore

$$\int_a^{a+h} f(x)\,dx \approx f(a)h \quad \text{and} \quad \lim_{h \to 0} \frac{1}{h}\int_a^{a+h} f(x)\,dx =$$

$$= \lim_{h \to 0} \frac{1}{h} f(a)h = f(a).$$

<u>Ex.</u> If X has a pdf f_X, find the pdf of Y, f_Y, where $Y = 2X$.

<u>Solution:</u> $F_Y(a) = P(Y \leq a) = P(2X \leq a) = P(X \leq \frac{a}{2})$

$= F_X\left(\frac{a}{2}\right)$

Hence $f_Y(a) = \frac{d}{da} F_Y(a) = \frac{d}{da} F_X\left(\frac{a}{2}\right) = \frac{1}{2} f_X\left(\frac{a}{2}\right)$

<u>Ex.</u> Suppose X has a pdf given by

$$f_X(x) = \begin{cases} \frac{1}{2} & 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

Find the pdf of $Y = \frac{1}{X}$

<u>Solution:</u> $0 < X < 2 \implies \frac{1}{2} < Y < \infty$

$F_Y(a) = P(Y < a) = P\left(\frac{1}{X} < a\right) = P\left(X > \frac{1}{a}\right)$

$$= 1 - P\left(X \leq \frac{1}{a}\right) = 1 - F_X\left(\frac{1}{a}\right)$$

Thus $f_Y(a) = \frac{d}{da} F_Y(a) = \frac{d}{da}\left(1 - F_X\left(\frac{1}{a}\right)\right)$

$$= \frac{1}{a^2} f_X\left(\frac{1}{a}\right)$$

Hence $f_Y(y) = \begin{cases} \frac{1}{2y^2} & \frac{1}{2} < y < \infty \\ 0 & \text{otherwise} \end{cases}$

<u>Ex.</u> X has pdf given by

$$f_X(x) = \begin{cases} e^{-x} & 0 < x < \infty \\ 0 & \text{otherwise} \end{cases}$$

Find the pdf of $Y = e^X$

<u>Solution:</u> $F_Y(a) = P(Y < a) = P(e^X < a) = P(X < \ln a)$

$$= F_X(\ln a)$$

Hence $f_Y(a) = \frac{d}{da} F_Y(a) = \frac{d}{da} F_X(\ln a) = \frac{1}{a} f_X(\ln a)$

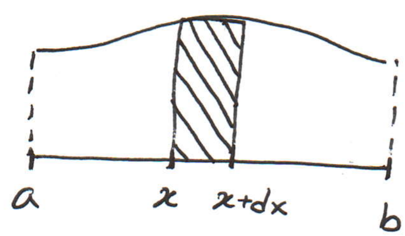
$$= \frac{1}{a} e^{-\ln a} = \frac{1}{a} \cdot \frac{1}{e^{\ln a}} = \frac{1}{a^2}$$

so $f_Y(y) = \begin{cases} \frac{1}{y^2} & -\infty < y < \infty \\ 0 & \text{otherwise} \end{cases}$

Expectation and Variance of Continuous Random Variables

For small dx $P(x < X < x+dx) \approx f(x)dx$. Thus

$$E[X] \approx \sum_{x} x f(x) dx \longmapsto \int_{-\infty}^{\infty} x f(x) dx.$$



In other words, we may view a continuous random variable as a discrete random variable whose probabilities are the bars $f(x)dx$, whose tips form the curve $f(x)$. In this format it is clear that the definition for expected value is carried forward to continuous random variables without change.

Ex. Find $E[X]$ when pdf of X is

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Solution: $E[X] = \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 2x^2 dx = \frac{2}{3} x^3 \Big|_0^1 =$

$= \frac{2}{3}.$

Ex. The density function of X is given by

$$f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $E[e^X]$

Solution: Let $Y = e^X$. We need to find pdf for Y.

$F_Y(a) = P(Y \leq a) = P(e^X \leq a) = P(X \leq \ln a) = F_X(\ln a)$

Hence $f_Y(a) = \frac{d}{da} F_Y(a) = \frac{d}{da} F_X(\ln a) = \frac{1}{a} f(\ln a)$.

Since $0 \leq x \leq 1$, $1 \leq e^x \leq e$. In other words

$$f_Y(y) = \begin{cases} \frac{1}{y} & 1 \leq y \leq e \\ 0 & \text{otherwise} \end{cases}$$

Thus $E[Y] = \int_1^e y \frac{1}{y} dy = e - 1$

Ex. The density function of X is given by

$$f(x) = \begin{cases} 3x^2 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find $E[\sqrt{X+1}]$

Solution: Let $Y = \sqrt{X+1}$. Since $0 \leq X \leq 1$,

$1 \leq Y \leq \sqrt{2}$.

$F_Y(a) = P(Y \leq a) = P(\sqrt{X+1} \leq a) = P(X \leq a^2-1)$

$= F_X(a^2-1)$.

Hence $f_Y(a) = \frac{d}{da} F_Y(a) = \frac{d}{da} F_X(a^2-1) = 2a f(a^2-1)$

It follows that the pdf of Y is

$$f_Y(y) = \begin{cases} 6y(y^2-1)^2 & 1 \leq y \leq \sqrt{2} \\ 0 & \text{otherwise} \end{cases}$$

$$E[Y] = \int_1^{\sqrt{2}} 6y^2(y^2-1)^2\,dy = \int_1^{\sqrt{2}} 6y^2(y^4-2y^2+1)\,dy$$

$$= \int_1^{\sqrt{2}} 6(y^6-2y^4+y^2)\,dy = 6\left(\frac{y^7}{7} - 2\frac{y^5}{5} + \frac{y^3}{3}\right)\Bigg|_1^{\sqrt{2}}$$

$$= \frac{132\sqrt{2} - 48}{105} \approx 1.320$$

The work of finding the pdf of $Y = g(X)$ can be rather tedius. Might we not work with the pdf of X instead?

<u>Ex.</u> Find $E[e^X]$ when pdf of X is

$$f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

<u>Solution:</u> $E[e^X] = \int_0^1 e^x \, dx = e-1$

<u>Ex.</u> Find $E[\sqrt{X^{\cancel{2}}+1}]$ when pdf of X is

$$f(x) = \begin{cases} 3x^2 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

<u>Solution:</u> $E[\sqrt{X+1}] = \int_0^1 \sqrt{x+1} \; 3x^2 \, dx$

let $u = \sqrt{x+1}$, $u^2 = x+1$, $x = u^2-1$, $dx = 2u\,du$

Hence $E[\sqrt{X+1}] = \int_1^{\sqrt{2}} 3u(u^2-1)^2 \cdot 2u \, du$

and the calculation follows in the same manner we solved this problem before.

<u>Proposition:</u> Let X be a continuous random variable with pdf $f(x)$. Let $Y = g(X)$ be continuous with $Y \geq 0$. Then $E[Y] = \int_{-\infty}^{\infty} g(x) f(x) \, dx$.

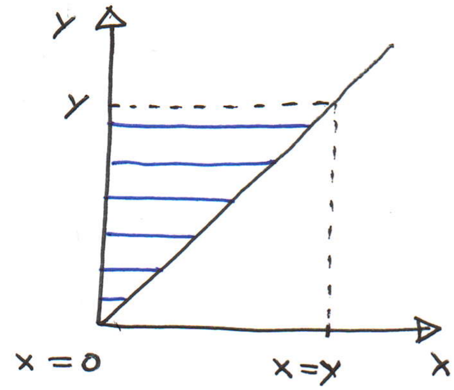
<u>Proof:</u> We may well argue that the result is an immediate consequence of the corresponding proposition in the discrete case; Merely regard X ~~as~~ a discrete r.v. with $P(X=x) = f(x)dx$, and Y as a discrete r.v. with $P(Y=y) = \sum_{x:\, g(x)=y} f(x)dx$. However, this is a good opportunity to

remind you to review ideas from multivariable calculus.

Let f_Y be the pdf of Y. By definition

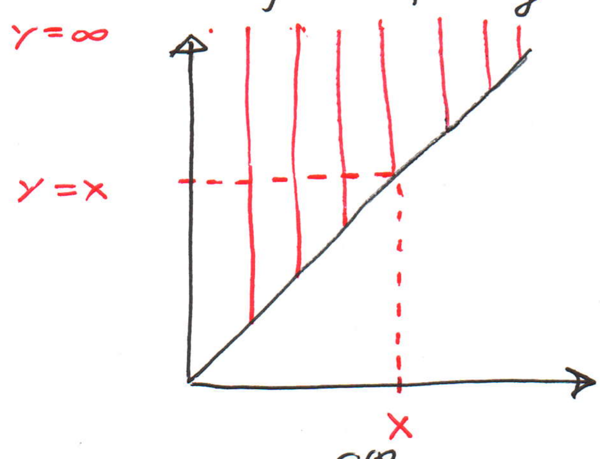
$$E[Y] = \int_0^\infty y\, f_Y(y)\,dy = \int_0^\infty \int_0^y f_Y(y)\,dx\,dy$$

This is a cake! In what baking pan is it made?



This is an infinite triangular pan and the cake is sliced by horizontal cuts (see my Calc. III notes).

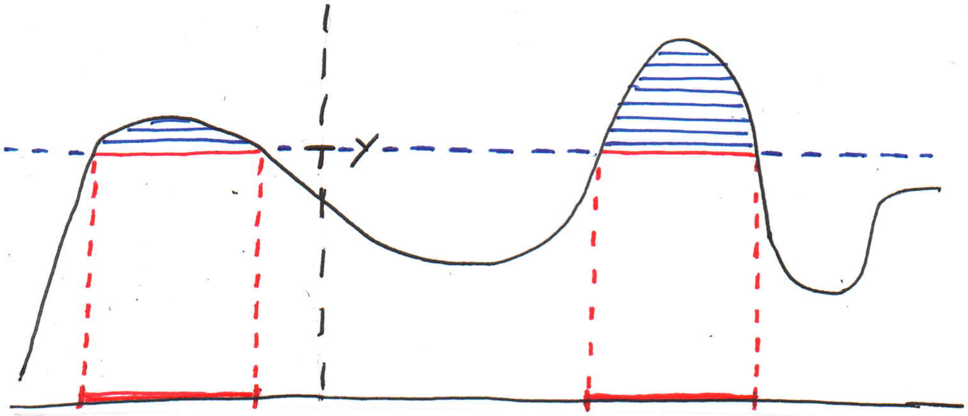
Changing order of integration, we get



$$\int_0^\infty \int_x^\infty f_Y(y)\,dy\,dx = \int_0^\infty P(Y > x)\,dx = \int_0^\infty P(Y > y)\,dy$$

$$= \int_0^\infty P(g(X) > y)\,dy = \int_0^\infty \int_{x:\, g(x) > y} f(x)\,dx\,dy$$

The last integral represents



$x: g(x) > y$

Thus $\displaystyle\int_0^\infty \int_{x:g(x)>y} f(x)\,dx\,dy = \int_{-\infty}^{\infty} \int_0^{g(x)} f(x)\,dy\,dx$

$$= \int_{-\infty}^{\infty} g(x) f(x)\,dx.$$

As it happens the result may be extended to arbitrary continuous functions $Y = g(X)$. The key is to define $g^+(x) = \max\{g(x), 0\}$ and $g^-(x) = \max\{-g(x), 0\}$. Then $g^+(x), g^-(x) \geq 0$ and $g(x) = g^+(x) - g^-(x)$. Using this hint see if you can prove the general Proposition

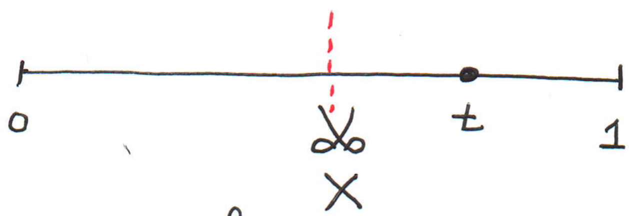
<u>Proposition:</u> If X is a continuous random variable with pdf $f(x)$, then for any real-valued function g,

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\,dx.$$

<u>Ex.</u> A stick of length 1 is split at a point X that is uniformly distributed over [0,1]. Determine the expected length of the piece that contains the point $0 \le t \le 1$.

<u>Solution:</u> The pdf of X is given by

$$f(x) = \begin{cases} 1 & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$



$$L_t(X) = \begin{cases} 1-X & \text{if } X \le t \\ X & \text{if } X > t \end{cases}$$

$$E[L_t(X)] = \int_0^t (1-x)\,dx + \int_t^1 x\,dx = t - \tfrac{1}{2}t^2 + \tfrac{1}{2}(1-t^2)$$

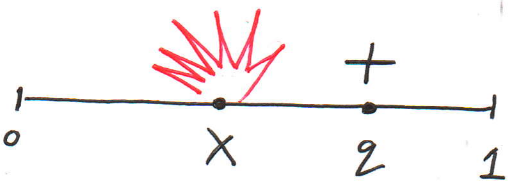
$= \frac{1}{2} + t(1-t)$ This is maximal when $t = \frac{1}{2}$

<u>Ex.</u> The place where an accident occurs is a random variable X with pdf

$$f(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

where should a hospital be placed to minimize expected distace from accident to hospital?

Solution:



$$D_q(X) = \begin{cases} q - X & \text{if } X \leq q \\ X - q & \text{if } X > q \end{cases}$$

$$E\left[D_q(X)\right] = \int_{-\infty}^{\infty} D_q(x) f(x)\,dx = \int_0^q (q-x)\cdot 2x\,dx +$$

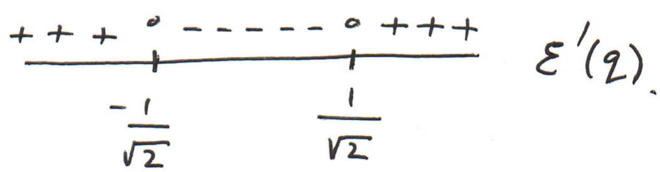
$$+ \int_q^1 (x-q)\, 2x\,dx = q\cdot q^2 - \frac{2}{3}q^3 - q + q^3 = \frac{2}{3} - q + \frac{2}{3}q^3$$

Setting $\mathcal{E}(q) = E\left[D_q(X)\right] = \frac{2}{3} - q + \frac{2}{3}q^3$

we see that $\mathcal{E}'(q) = 0$ when $2q^2 - 1 = 0$ or when

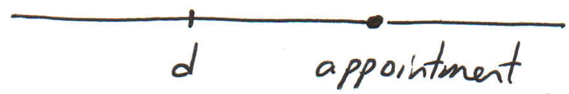
$q = \frac{1}{\sqrt{2}}$ Thus we should build hospital at position

$q = \frac{1}{\sqrt{2}}$.



Ex. If you're s minutes early to appointment you incur a cost of cs. If you're s minutes late, the cost is ks. Travel time to destination is a continuous random variable T with pdf f. At what time should you depart to minimize your expected cost?

<u>Solution:</u> Let d be the time of departure before appointment.



$$\text{Let } C_d(t) = \begin{cases} c(d-t) & \text{if } t \leq d \\ k(t-d) & \text{if } t > d \end{cases}$$

$$E[C_d(T)] = \int_0^d c(d-t)f(t)\,dt + \int_d^\infty k(t-d)f(t)\,dt = \mathcal{E}(d)$$

$$\mathcal{E}(d) = cd\int_0^d f(t)\,dt - \int_0^d ctf(t)\,dt + \int_d^\infty kt\,f(t)\,dt - kd\int_d^\infty f(t)\,dt$$

$$\mathcal{E}'(d) = c\,F(d) + cdf(d) - cdf(d) - kd\,f(d) - k(1-F(d))$$

$$+\,kdf(d) = (c+k)F(d) - k = 0, \quad F(d) = \frac{k}{c+k}$$

$$\begin{array}{c} - - - \; 0 \; + + + \\ \hline \dfrac{k}{c+d} \end{array} \quad \mathcal{E}'(d)$$

In particular, depart at a moment d such that the probability of taking d or less units of time to arrive at appointment is $\frac{k}{c+k}$.

Corollary: If a and b are constants then

$$E[aX+b] = aE[X]+b.$$

proof: $E[aX+b] = \int_{-\infty}^{\infty}(ax+b)f(x)dx =$

$$= a\int_{-\infty}^{\infty} x f(x)dx + b\int_{-\infty}^{\infty} f(x)\,dx = aE[X] + b\cdot 1.$$

Note that variance is defined in the same way for continuous random variables as for discrete ones. Namely $Var(X) = E[(X-\mu)^2]$ where $\mu = E[X]$.

Clearly $Var(X) = E[X^2] - (E[X])^2$ and

$$Var(aX+b) = a^2\,Var(X).$$

Ex. Let X be continuous random variable with pdf defined by

$$f(x) = \begin{cases} 2x & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Compute $E[X]$ and $Var(X)$.

Solution: $E[X] = \int_0^1 2x^2 dx = \frac{2}{3}$

$$E[X^2] = \int_0^1 2x^3 dx = \frac{2}{4} = \frac{1}{2}$$

$$Var(X) = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{9-8}{18} = \frac{1}{18}.$$