

8.5a

probability

(1)

Weighted averages are familiar in the discrete case.

Suppose N students take a test with grades X_1, \dots, X_N between 0 and 100, but there are only $n \leq N$ distinct grades x_1, \dots, x_n with each such grade repeated $f(x_1), \dots, f(x_n)$ times, where clearly $\sum_{i=1}^n f(x_i) = N$ must be the total number of students.

Then the average test grade can be evaluated by grouping together all matching test grades:

$$\langle x \rangle = \frac{\sum_{i=1}^N X_i}{N} = \frac{f(x_1)x_1 + \dots + f(x_n)x_n}{N} = \sum_{i=1}^n x_i \frac{f(x_i)}{N} = \sum_{i=1}^n x_i p(x_i)$$

where $p(x_i)$ is the fraction of students getting the test grade x_i

$$\text{and } \sum_{i=1}^n p(x_i) = \sum_{i=1}^n \frac{f(x_i)}{N} = \frac{1}{N} \sum_{i=1}^n f(x_i) = \frac{1}{N}(N) = 1$$

they must add up to one.

We just weight each grade by the fraction of students getting that grade and add up all these products.

If we picked a student "at random", these weighting factors would give the probabilities that each student would receive a given grade.

We have n discrete values of the "random variable" x (testgrade) on the interval $0 \leq x \leq 100$, each with an associated probability which is distributed among those grades. This "distribution" is the set of proper fractions which add up to 1.

Summary: $\langle x \rangle = \sum_{i=1}^n x_i p(x_i), \quad \sum_{i=1}^n p(x_i) = 1, \quad 0 \leq p(x_i) \leq 1$

The Riemann definition of an integral helps us transfer this to a "continuous distribution".

Remark: The probability that a student receives a grade belonging to a subset of these possible grades is just the sum of the probabilities of that subset of grades

8.5a

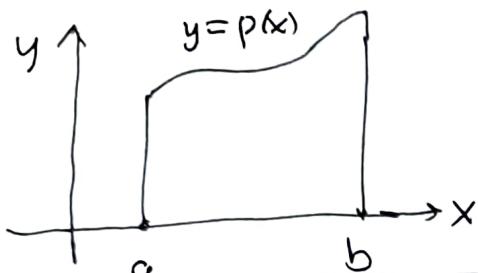
probability

(2)

Continuous probability distribution

Suppose we have a variable x on an interval $a \leq x \leq b$
 which are possible values of some quantity.

A "distribution" of such values is a non-negative function



$$0 \leq p(x) \leq 1$$

$$\int_a^b p(x) dx = 1 \quad (\leftrightarrow \sum_{i=1}^n p(x_i) = 1)$$

discrete case:

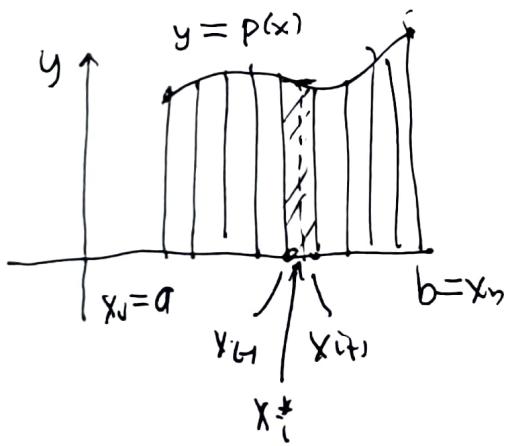
The "average" or expected value of x is:

$$\langle x \rangle = \int_a^b x p(x) dx \quad (\leftrightarrow \sum_{i=1}^n x_i p(x_i))$$

discrete case:

"probability distribution function"
 = weighting factor

Riemann shows how this comes from a limiting process of a discrete distribution



We get a set of sampled values

$$x_i^*, i = 1, \dots, n$$

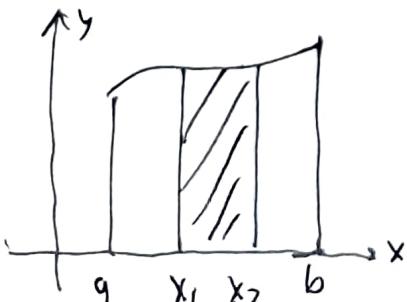
we weight each value by the area under the curve (approximately)

$$x_i^* p(x_i^*) \Delta x \quad \begin{array}{l} \text{note } \sum_{i=1}^n p(x_i^*) \Delta x \\ \approx \int_a^b p(x) dx = 1 \end{array}$$

discrete probability distribution

Then we average these values multiplied by their weighting factors

$$\langle x \rangle \approx \sum_{i=1}^n x_i^* p(x_i^*) \Delta x \longrightarrow \int_a^b x p(x) dx \equiv \langle x \rangle$$



The probability that x assumes a value in an interval is the area under the curve

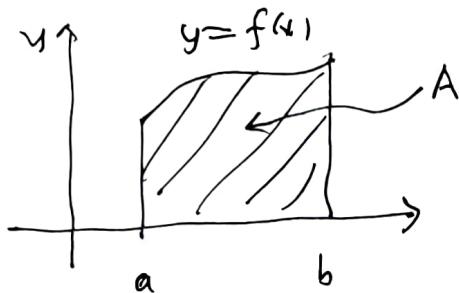
$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} p(x) dx$$

8.5a

probability

(3)

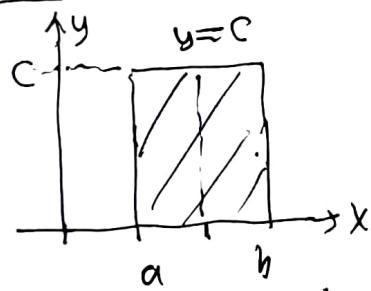
Example: Given any function $f(x) \geq 0$ on an interval (closed or semi-infinite or infinite), if its integral is finite, one can divide the function by that integral to get a probability distribution on that interval.



$$\int_a^b f(x) dx = A \quad (\text{area under curve})$$

$$\text{Then define } p(x) = \frac{f(x)}{A}$$

$$\begin{aligned} \int_a^b p(x) dx &= \int_a^b \frac{f(x)}{A} dx = \frac{1}{A} \int_a^b f(x) dx \\ &= \frac{1}{A}(A) = 1. \end{aligned}$$

uniform distribution

$$\int_a^b C dx = C(b-a)$$

$$p(x) = \frac{C}{C(b-a)} = \frac{1}{b-a}$$

of course
 $(b-a)\left(\frac{1}{b-a}\right) = 1$
rectangle!

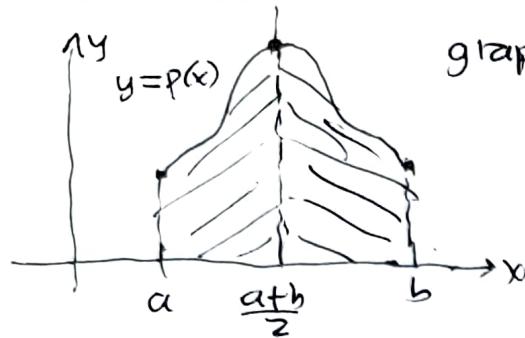
$$\begin{aligned} \mu = \langle x \rangle &= \int_a^b x \left(\frac{1}{b-a} \right) dx = \frac{x^2}{2(b-a)} \Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} \\ &= \frac{1}{2}(a+b) \quad \text{average of endpoints.} \end{aligned}$$

↑
Greek letter mu traditionally used for average or expected value

8.5a

Probability

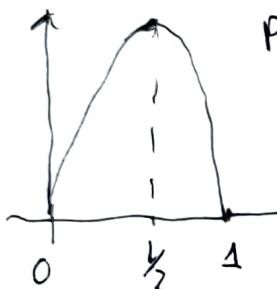
(4)

symmetric distribution

symmetric about center

graph reflects across center of interval.
equal area on either side of midpoint.
half probability allotted to each side.

example start with $f(x) = x(1-x)$ $0 \leq x \leq 1$

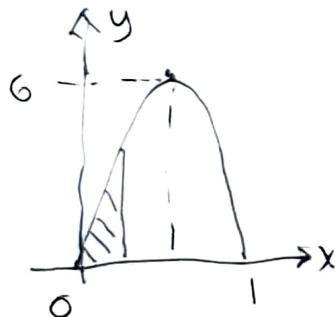


parabola

$$\begin{aligned} A &= \int_0^1 x(1-x) dx = \int_0^1 x - x^2 dx \\ &= \left[\frac{x^2}{2} - \frac{x^3}{3} \right]_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{3-2}{6} = \frac{1}{6} \end{aligned}$$

so define $P(x) = 6x(1-x)$

$$\begin{aligned} \langle x \rangle &= \int_0^1 x \cdot 6x(1-x) dx = 6 \int_0^1 x^2 - x^3 dx = 6 \left(\frac{x^3}{3} - \frac{x^4}{4} \right) \Big|_0^1 \\ &= 6 \left(\frac{1}{3} - \frac{1}{4} \right) = 6 \left(\frac{4-3}{12} \right) = \frac{1}{2} \quad \checkmark \text{ midpoint.} \end{aligned}$$



$$\begin{aligned} P(0 \leq x \leq \frac{1}{4}) &= \int_0^{\frac{1}{4}} 6x(1-x) dx = 6 \int_0^{\frac{1}{4}} x - x^2 dx \\ &= 6 \left(\frac{x^2}{2} - \frac{x^3}{3} \right) \Big|_0^{\frac{1}{4}} = 6 \left(\frac{1}{2} \left(\frac{1}{16} \right) - \frac{1}{3} \left(\frac{1}{64} \right) \right) \\ &= \frac{6}{32} \left(1 - \frac{1}{3 \cdot 2} \right) = \frac{6}{32} \left(\frac{5}{6} \right) = \frac{5}{32} \approx 0.156 \end{aligned}$$

 $\sim 16\%$ probability

(nearest percent good enough usually)
not rocket science!

BONUS

middle half!

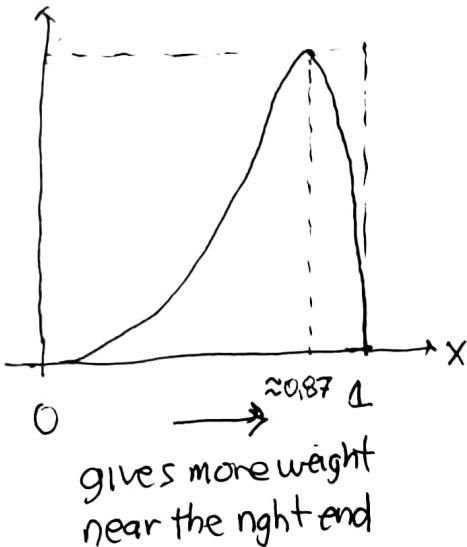
$$1 - 2(0.156) \approx 0.6875 \approx 0.69$$

$\sim 69\%$ probability,
about $2/3$ roughly.

8.5a

Probability

(5)

Asymmetric distributionstart with $f(x) = x^3 \sqrt{1-x^2}$

$$A = \int_0^1 x^2 (1-x^2)^{1/2} \cdot x \, dx$$

$\overbrace{1-u}^{\text{u}}$ \overbrace{u}^{u} $\overbrace{-\frac{1}{2} du}^{\text{d}u}$

} easily done by u-sub

$$\text{Maple} = \frac{2}{15}$$

$$p(x) = \frac{15}{2} x^3 (1-x^2)^{1/2}$$

But we cannot ourselves evaluate the expected value—we need Maple:

$$\mu \equiv \langle x \rangle = \int_0^1 x \cdot \frac{15}{2} x^3 (1-x^2)^{1/2} \, dx \approx 0.736$$

~ 0.74

What is the probability x assumes a value larger than $1/2$?

$$P\left(\frac{1}{2} \leq x \leq 1\right) = \int_{1/2}^1 \frac{15}{2} x^3 (1-x^2)^{1/2} \, dx \approx 0.89$$

Maple

This we can do by hand BUT
the important thing is knowing how to setup the integral & what it represents

8.5a

Probability

(6)

Semi-infinite distribution

Suppose you wait for something to happen starting at time $t=0$. There is no largest value of time to delimit the waiting time, so it requires $0 \leq t < \infty$ BUT clearly very large times should contribute very little to the total probability that something happens.

example Poisson = decaying exponential distribution

start with $f(x) = e^{-kx}$, $k > 0$

$$A = \int_0^\infty e^{-kx} dx = \left[\frac{e^{-kx}}{-k} \right]_0^\infty = \lim_{t \rightarrow \infty} -\frac{1}{k} e^{-kt} \Big|_0^t$$

$$= \lim_{t \rightarrow \infty} -\frac{1}{k} e^{-kt} + \frac{1}{k} e^0 = \frac{1}{k}$$

$$\text{so } p(x) = \frac{e^{-kx}}{k} = k e^{-kx}$$

Calculate

$$\mu = \langle x \rangle = \int_0^\infty kx e^{-kx} dx = \left[-\frac{e^{-kx}(kx+1)}{k} \right]_0^\infty$$

$$= \lim_{t \rightarrow \infty} -\frac{e^{-kt}(kt+1)}{k} \Big|_0^t = \lim_{t \rightarrow \infty} \left(\frac{e^{-kt}(kt+1)}{k} + e^0 \left(\frac{1}{k} \right) \right)$$

$$= \lim_{t \rightarrow \infty} \left(-\frac{kt+1}{k e^{kt}} + \frac{1}{k} \right) = \frac{1}{k} \quad \text{so } p(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}}$$

↑
exp's beat polys

$$k = \frac{1}{\mu} \quad = \frac{1}{\mu} e^{-x/\mu}$$

now expressed in terms of meaningful parameter instead of k .

8.5a) Probability

(7)

Textbook example 4 (WORD PROBLEM)

Suppose avg time for customer service to pick up is 5 minutes.

Assuming the $\mu=5$ exponential distribution:

a) What is $P(0 \leq t \leq 1)$ probability that call is answered in first minute

b) What is $P(t \geq 5)$ probability that customer waits more than 5 min.

Soln

$$\text{a) } P(0 \leq t \leq 1) = \int_0^1 \frac{1}{5} e^{-t/5} dt = \frac{1}{5} \left[e^{-t/5} \right]_0^1 = -\left(e^{-1/5} - e^0 \right) = 1 - e^{-1/5} \approx 0.181 \\ \approx 18\%$$

"About 18% of customer calls are answered in the first minute."

$$\text{b) } P(t \geq 5) = \int_5^\infty \frac{1}{5} e^{-t/5} dt = \frac{1}{5} \left[e^{-t/5} \right]_5^\infty = \lim_{t \rightarrow \infty} -e^{-t/5} \Big|_5^\infty = \lim_{t \rightarrow \infty} (-e^{-t/5} + e^{-5/5}) = e^{-1} \\ \approx 0.368 \sim 37\%$$

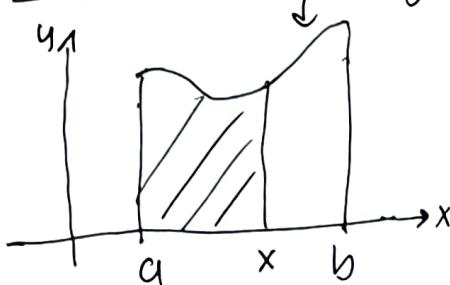
"About 37% of customers wait more than 5 minutes before their calls are answered."

8.5a

Probability

(8)

Median value?


 $y = p(x)$ - "probability distribution function" = PDF

area up to x = probability that this random variable x assumes a value between a and x

$$\text{CDF}(x) = \underbrace{\int_a^x p(t) dt}_{\text{just the particular antiderivative of } p(x) \text{ which vanishes at } x=a.} \quad \begin{array}{l} \text{"cumulative distribution function"} \\ = \text{area accumulation function starting from the left endpoint.} \end{array}$$

This increases from 0 to 1 as one moves across the probability interval.

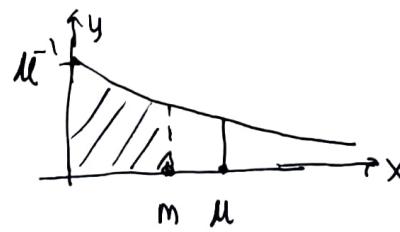
When it reaches $\frac{1}{2}$, we are at the median value of x .

median $\Leftrightarrow \int_a^x p(t) dt = \frac{1}{2}$, solve for x .

example $p(x) = \mu^{-1} e^{-x/\mu}, \quad x \geq 0, \quad \mu = \langle x \rangle = \text{mean value}$

$$\frac{1}{2} = \int_0^x \mu^{-1} e^{-t/\mu} dt = \mu^{-1} \left[e^{-t/\mu} \right]_0^x = -e^{-x/\mu} + 1$$

solve: $1 - \frac{1}{2} = e^{-x/\mu}$
 $\frac{1}{2} = e^{-x/\mu} \rightarrow 2 = e^{x/\mu} \rightarrow \ln 2 = \frac{x}{\mu} \rightarrow x = \mu \ln 2 = \mu \cdot 0.69 = M$



In a call center wait time application, if we average the wait time for 100 calls, we get the mean value μ .

on the other hand there is a 50% chance the callee will have to wait up to M minutes. (and a 50% chance the callee will have no wait longer.)

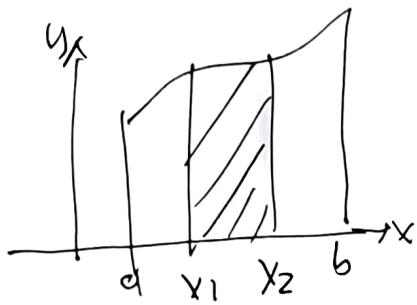
8.5a

Probability

(9)

$$P(X_1 \leq X \leq X_2) = \int_{X_1}^{X_2} P(x) dx = \underbrace{CDF(x)}_{\text{any antiderivative}} \Big|_{X_1}^{X_2} = CDF(X_2) - CDF(X_1)$$

any
antiderivative
will do



If $CDF(x)$ is tabulated, its differences give desired probabilities.

Before IT saved us, tabulated values in books were consulted.