Probability Distributions

- First imagine that a variable x takes a set of discrete outcomes x_i, e.g. a loaded dice with probability P_i $i = 1 \dots N$, with $N = 6$ for loaded dice

Then
$$\sum_i P_i = 1$$

$$\langle x \rangle = \sum_i x_i P_i \qquad \text{also } \bar{x}$$

$$\langle x^2 \rangle = \sum_i x_i^2 P_i \qquad \text{also } \overline{x^2}$$

The deviation from the mean is

$$\delta x \equiv x - \langle x \rangle$$

The mean deviation is

$$\langle \delta x \rangle = \langle x - \underbrace{\langle x \rangle}_{\text{const}} \rangle = \langle x \rangle - \langle x \rangle = 0$$

- The mean squared deviation, or variance

$$\sigma_x^2 \equiv \langle \delta x^2 \rangle = \langle (x - \langle x \rangle)^2 \rangle$$

std. deviation squared

$$= \langle x^2 - 2x\underbrace{\langle x \rangle}_{\text{const}} + \underbrace{\langle x \rangle^2}_{\text{const}^2} \rangle$$

$$= \langle x^2 \rangle - 2\langle x \rangle \langle x \rangle + \langle x \rangle^2$$

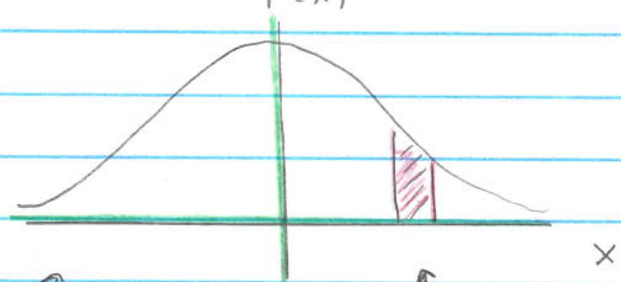
$$\boxed{\sigma_x^2 = \langle x^2 \rangle - \langle x \rangle^2}$$

Continuous variables:

- The book (and other books) call it $P(x)$ for the <u>probability density</u>.

$d\mathcal{P} = P(x)\,dx$, probability to find x between x and $x + dx$

probabity



$$\frac{d\mathcal{P}}{dx} = P(x)$$

$P(x)$ is the probability per x

Gaussian Distribution

$d\mathcal{P}$ = probability to be in this bin

Example: The Gaussian Distribution / Bell Curve

- $P(x) = N e^{-x^2/2\sigma^2}$
- Find N, $\langle x \rangle$, $\langle \delta x^2 \rangle$

Ok

$$1 = \sum d\mathcal{P} = \int P(x)\,dx$$

$$1 = \underbrace{\int_{-\infty}^{\infty} N e^{-x^2/2\sigma^2}\,dx}_{\text{Integral you should know}}$$

- If you dont know how to do this integral go read Appendix C2 through Eq. C6

$$1 = N \cdot (2\pi\sigma^2)^{1/2} \implies N = \frac{1}{\sqrt{2\pi\sigma^2}}$$

Thus we find the normalized Gaussian

$$\boxed{P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}}$$

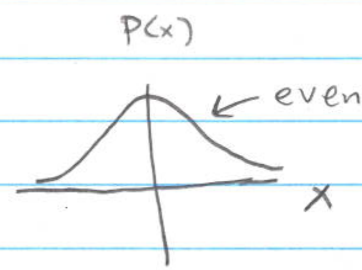
note the units are $1/$distance since σ has units distance

try to remember this

Now what about $\langle x \rangle$

$$\langle x \rangle = \int_{-\infty}^{\infty} x \; P(x)\, dx = 0$$

odd $\times$ even $=$ odd



- Then we want $\langle x^2 \rangle$

$$\langle x^2 \rangle = \int_{-\infty}^{\infty} dx \; x^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2} = \sigma^2$$

There is a trick to doing this integral (and many like this) called a generating function which we will describe in homework.

Note that

$\langle x^2 \rangle$ has units (distance)^2 as it should.

Also

$\sigma_x^2 = \langle x^2 \rangle - \cancel{\langle x \rangle^2}^{\,0} = \sigma^2$, justifying the name.

Change of Variables

- Given a probability distribution $d\mathcal{P} = P_x(x)\,dx$. and a change of variables $u(x)$ or $x(u)$. Then

$$\boxed{d\mathcal{P} = P_x(x)\,dx = P_x(x(u))\left|\frac{dx}{du}\right| du}$$

Thus

$$\boxed{P_u(u) = P_x(x(u))\left|\frac{dx}{du}\right|}$$

We have assumed that the map is one-to-one.

Why absolute value?

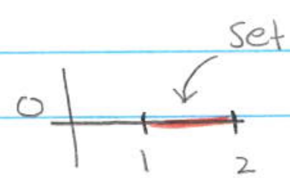
- Short answer: We are asking wether u is in a bin of width du and not wether du is increasing or decreasing. Ther probability is positive after all.

- Long answer:

You can always use the absolute value in change of variables while integrating, provided you regard the integral as over a region $[a,b]$ rather than an integral from a to b, ie. an un oriented integral rather than an oriented one.

Change of Variables in Integration

$$I = \int_1^2 x^2\,dx$$

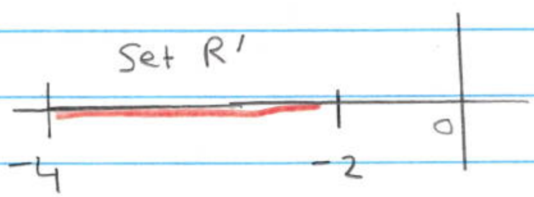


- under change of variables $u = -2x$ or $x = -u/2$ $dx = -\frac{du}{2}$

$$I = \int_{-2}^{-4} \left(\frac{-u}{2}\right)^2 \left(-\frac{1}{2}\,du\right) \qquad dx/du = -1/2$$

flip limits
$$= \int_{-4}^{-2} \left(\frac{-u}{2}\right)^2 \frac{1}{2}\,du \qquad \left|\frac{dx}{du}\right| = \frac{1}{2}$$

- So we are integrating over a new set (the mapped region R')



So the change of variables can alway be written

$$\boxed{I = \int_R dx\, f(x) = \int_{R'} du \left|\frac{dx}{du}\right| f(x(u))}$$

mapped region of integration

- This approach to integration, which ignores the orientation of the region (but with $\left|\frac{dx}{du}\right|$) is always valid. In fact there are sophisticated contexts where the absolute value must be used.