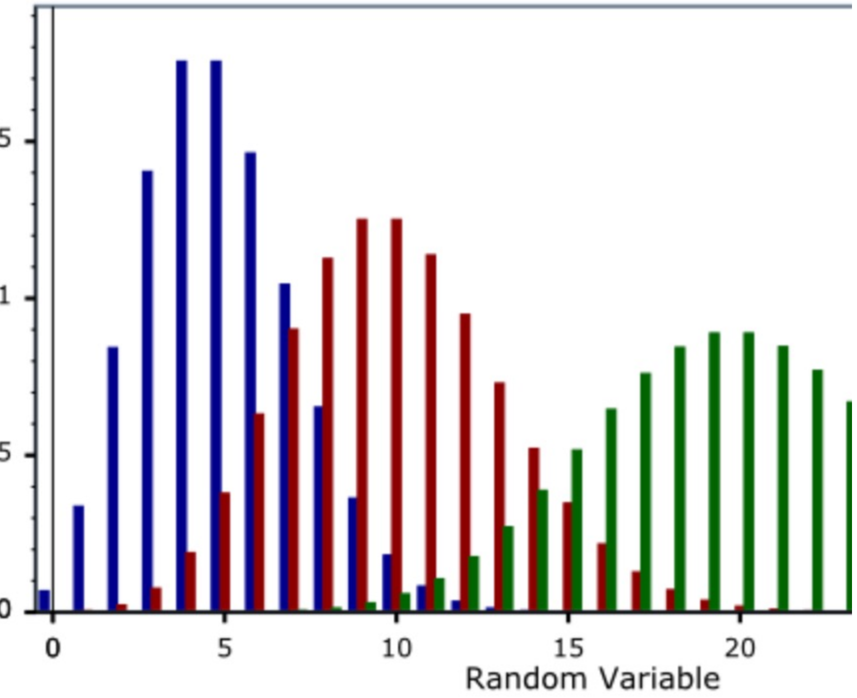
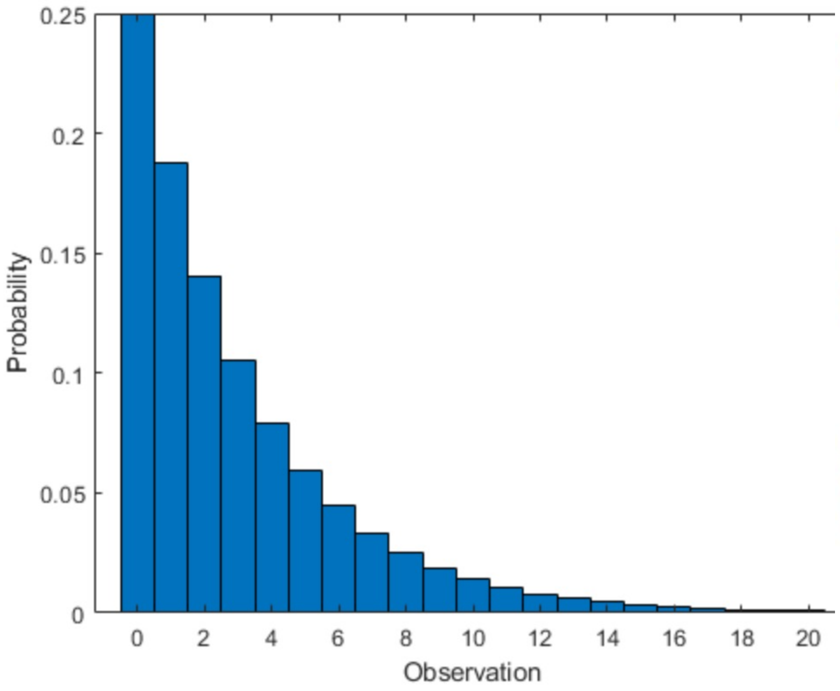
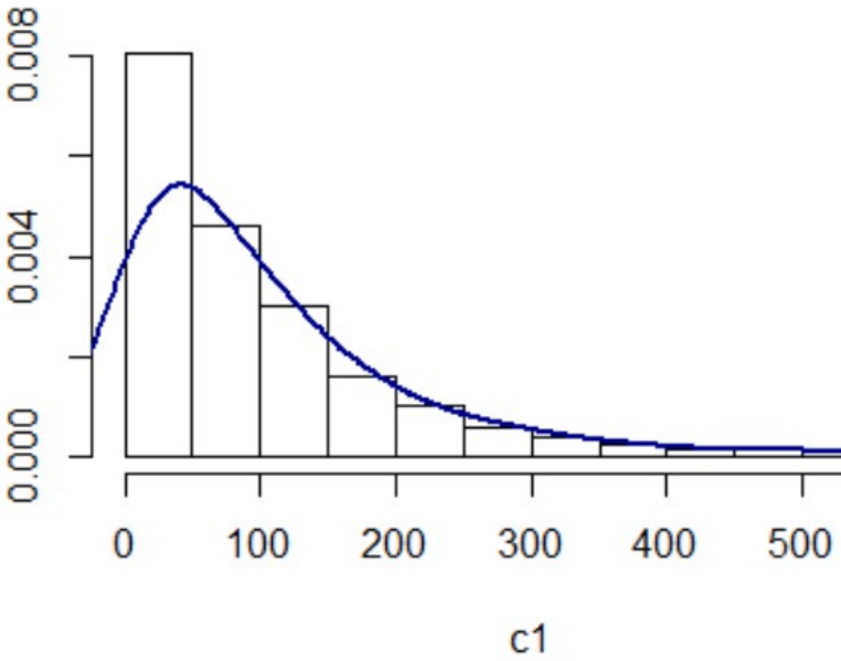
CHAPTER SEVEN

Expectation and Variance







- Random variables contain LOTS of information concerning the experiments that are modeled by them. How can this complication be simplified?

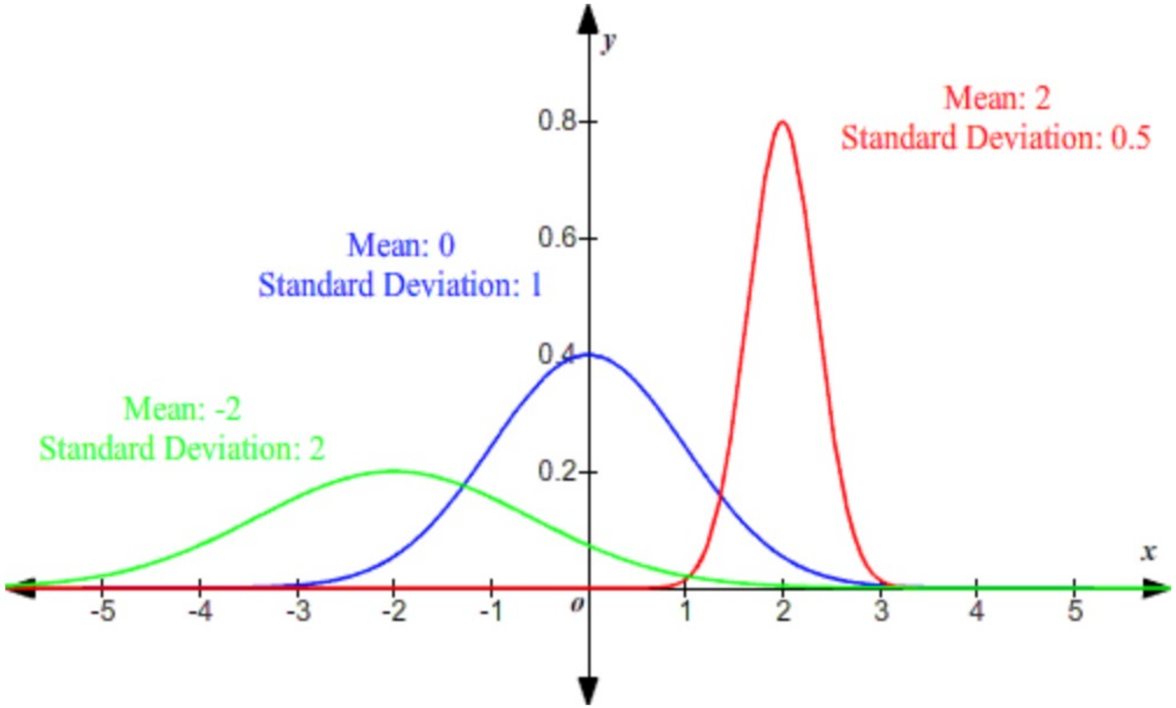
- Can we summarize each different random variable with one number?

The single number that might best describe a distribution would be its mean,

aka its expectation,

aka its expected value.

A second-best option would be its variance: the measure of spread of the distribution of the random variable.



- Example:
- An oil company is doing exploration drilling.

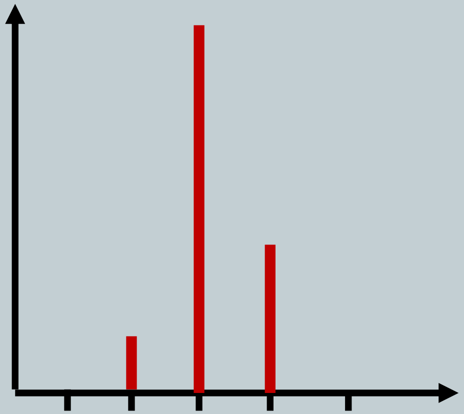
  They are using drill bits of type A.

Type A drill bits are known to last 2 hours 10% of the time, they last 3 hours 70% of the time, and they last 4 hours the remaining 20% of the time.

- Each time a bit wears out, it is replaced with another type A bit.
- The company is currently exploring an area and has 10 type A bits with them.

- How many hours would you estimate the crew can expect to explore before the bits are used up?

Hours	Probability
2	.10
3	.70
4	.20



This information corresponds to a random variable X whose distribution is given by :

$P(X = 2) = 0.1 \qquad P(X = 3) = 0.7 \qquad P(X = 4) = 0.2$

The expected number of exploration hours is best answered with the '*expected value*':

$E(X) = (0.1) \cdot 2 + (0.7) \cdot 3 + (0.2) \cdot 4 = 3.1$ hours per bit

Now, considering that the crew is carrying 10 bits:

$10 \cdot (3.1) =$ 31 hours of exploration time can be expected.

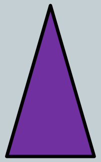
This drilling random variable X has a distribution given by :

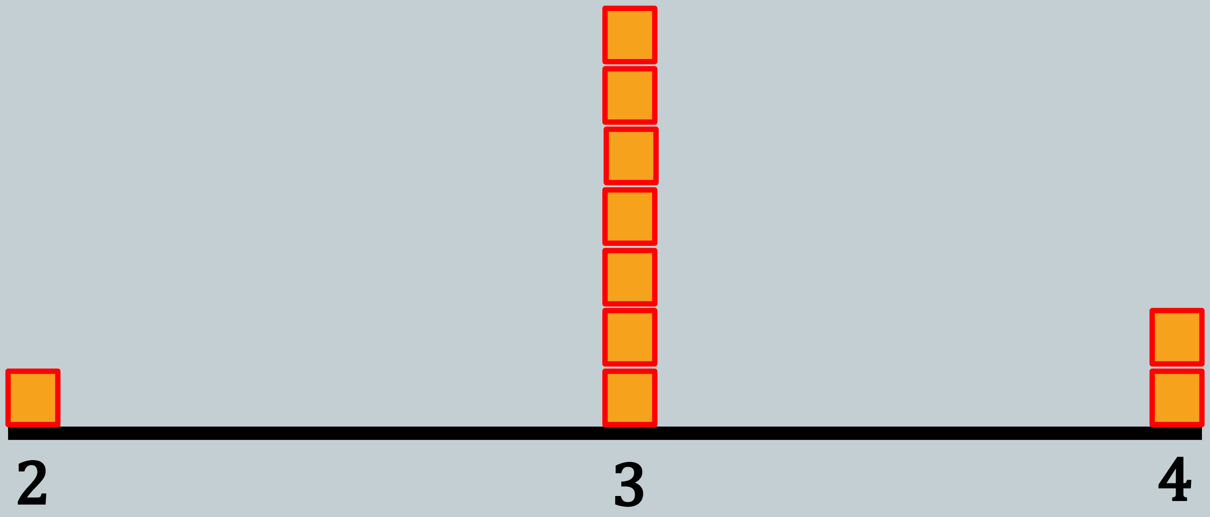
$P(X = 2) = 0.1 \qquad P(X = 3) = 0.7 \qquad P(X = 4) = 0.2$

The expected number of exploration hours per drill bit is :

$$E(X) = \sum_i a_i \, p(a_i) = (0.1) \cdot 2 + (0.7) \cdot 3 + (0.2) \cdot 4 = 3.1$$

Graphically, where would you put the fulcrum in order to create balance?





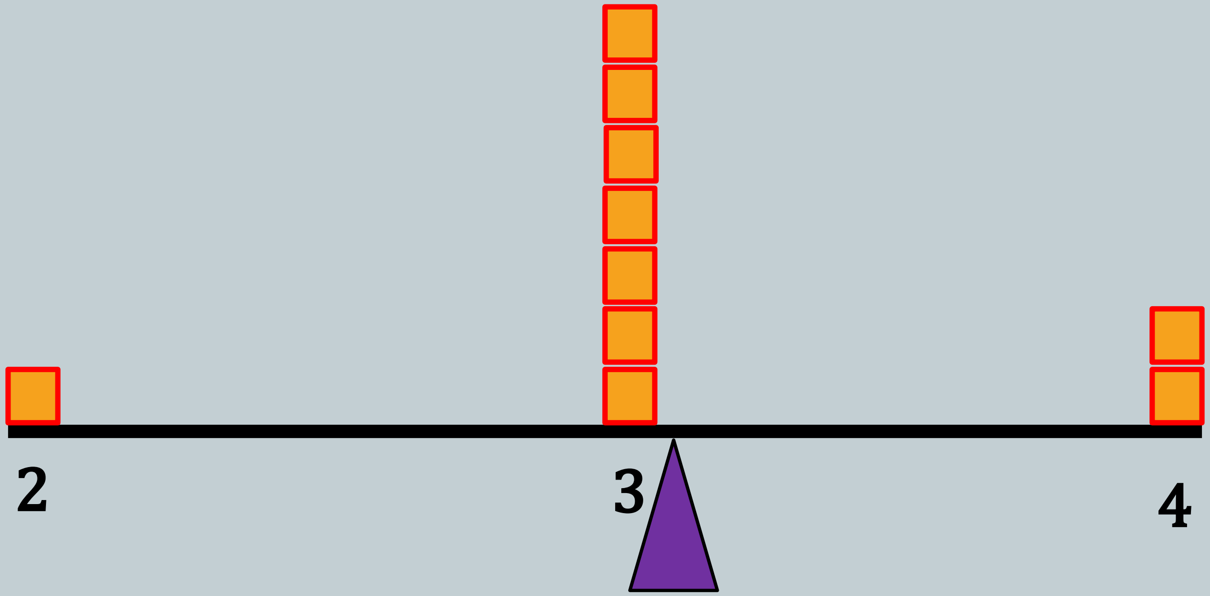
Random variable X has a distribution given by :

$P(X = 2) = 0.1$ $\quad P(X = 3) = 0.7$ $\quad P(X = 4) = 0.2$

'*Expected value*':

$E(X) = \sum_i a_i \, p(a_i) = (0.1) \cdot 2 + (0.7) \cdot 3 + (0.2) \cdot 4 = 3.1$

Graphically, where would you put the fulcrum in order to create balance?



Since some type A bits have been known to last only 2 hours and some have been known to last up to 4 hours, it might happen that the crew could be drilling for $10 \cdot 4 = 40$ hours, or they could be unlucky and only be able to drill for $10 \cdot 2 = 20$ hours.

However, 31 is the 'correct' answer in the sense that for a large number, n, of drill bits the total running time will be around $[n \cdot (3.1)]$ hours with high probability.

In General:

The Expectation of a discrete random variable X taking the values $a_1, a_2, \ldots$ and with probability mass function p is the number:

$$E[X] = \sum_i a_i P(X = a_i) = \sum_i a_i p(a_i)$$

$E[X]$ is called the expected value of X, or the mean of X.

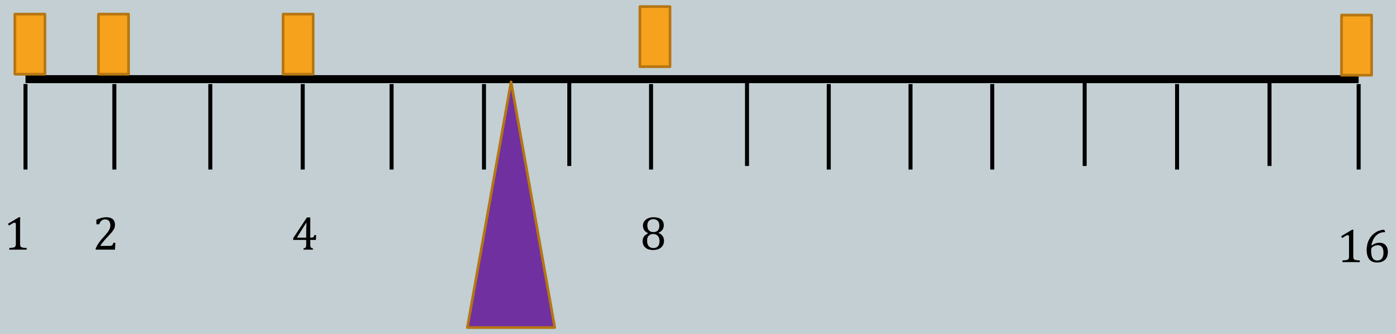
Example: Let X be the discrete random variable that takes the values: $1, 2, 4, 8, 16$ each with probability $\frac{1}{5}$.

What is $E[X]$?

Example: Let X be the discrete random variable that takes the values: $1, 2, 4, 8, 16$ each with probability $\frac{1}{5}$.

What is $E[X]$?

ANS: $E[X] = \sum_i a_i P(X = a_i) = 1 \cdot \frac{1}{5} + 2 \cdot \frac{1}{5} + 4 \cdot \frac{1}{5} + 8 \cdot \frac{1}{5} + 16 \cdot \frac{1}{5} = \frac{31}{5} = 6.2$



If the expected value for a discrete random value is given below, then what would you think the expected value for a continuous random variable would be?

The Expectation of a discrete random variable X taking the values $a_1, a_2, \ldots$ and with probability mass function p is the number:

$$E[X] = \sum_i a_i P(X = a_i) = \sum_i a_i p(a_i)$$

$E[X]$ is called the expected value of X, or the mean of X.

Expectation for a discrete random variable.

The Expectation of a discrete random variable X taking the values $a_1, a_2, \ldots$ and with probability mass function p is the number:

$$E[X] = \sum_i a_i\, P(X = a_i) = \sum_i a_i\, p(a_i)$$

$E[X]$ is called the expected value of X, or the mean of X.

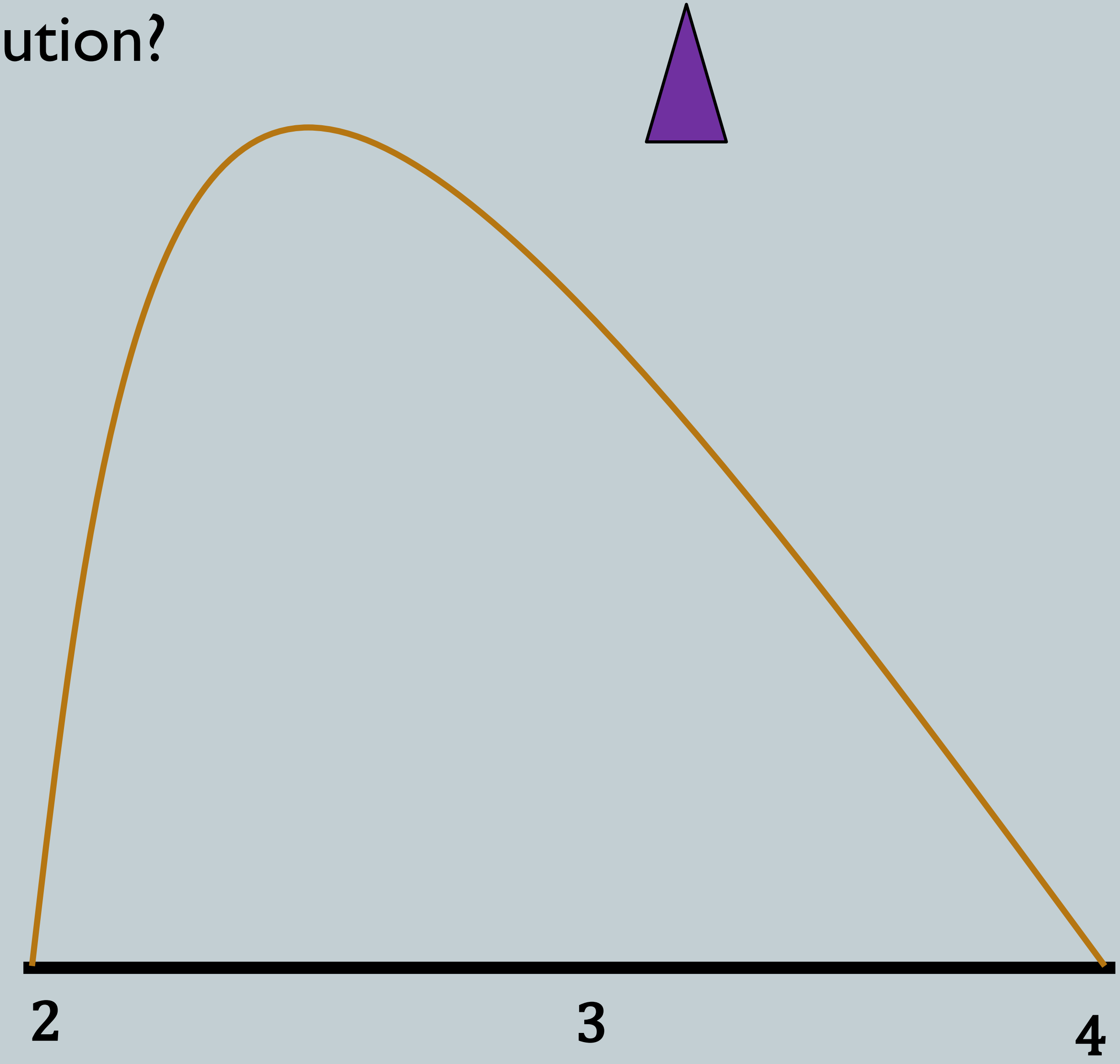
Expectation for a continuous random variable

The Expectation of a continuous random variable X with probability density function f is the number:

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

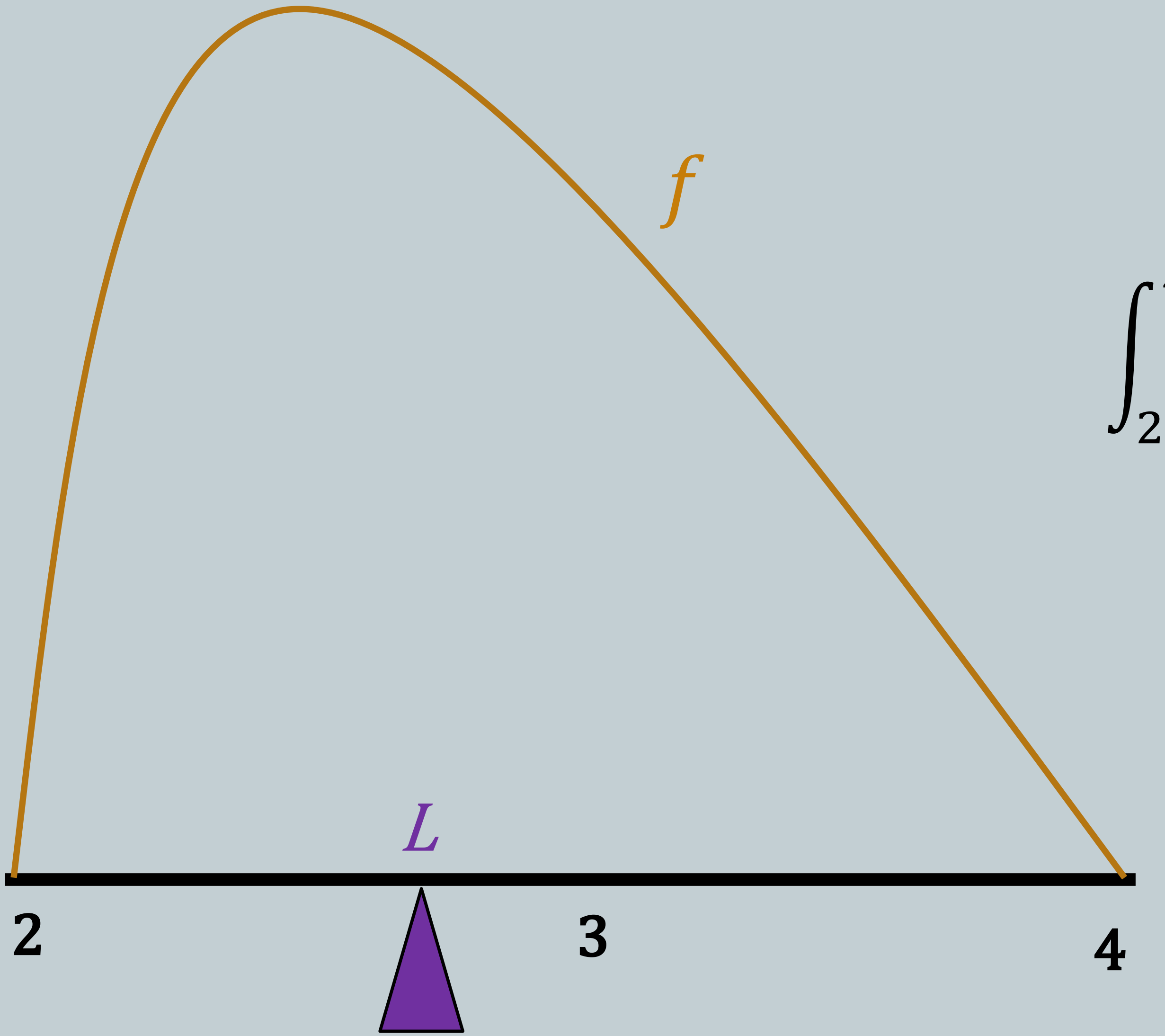
$E[X]$ is called the expected value of X, or the mean of X.

Graphically, where would you put the fulcrum in order to create balance for this continuous distribution?



Graphically, where would you put the fulcrum in order to create balance?

The fulcrum location is the expected value.



$$\int_2^4 x f(x)\, dx = \boldsymbol{L}$$

Suppose the odds of winning a lottery are 1 in 10,000. (PB is 1 in 293M, 5.6M yrs.)
You buy a lotto ticket every week.
Then the expected number of weeks you would have to buy tickets before you win would be 10,000 weeks (around 200 years).

If X is the random variable 'number of weeks till you win', then $E[X] = \frac{1}{10,000}$

$$P(X=1) = p$$
$$P(X=2) = (1-p)p$$
$$P(X=3) = (1-p)^2 p$$
$$P(X=4) = (1-p)^3 p$$
$$\vdots$$
$$P(X=n) = (1-p)^{n-1} p$$

Recall $E[X] = \sum_i a_i P(X = a_i) = \sum_i a_i\, p(a_i)$
Therefore, for this lotto example:

$$E[X] = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \frac{1}{p}$$

notice $\sum_i a_i\, p(a_i) = 1 \cdot p + 2 \cdot (1-p)p + 3 \cdot (1-p)^2 p + 4 \cdot (1-p)^3 p + \cdots$

The expectation of a Geometric Distribution:

Let X have a geometric distribution with parameter p then $E[X] = \frac{1}{p}$

$$\boxed{E[X] = \sum_i a_i P(X = a_i)} \quad \Rightarrow \quad \boxed{E[X] = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \frac{1}{p}}$$

The expectation of an exponential distribution:

Let X be an exponential distribution with parameter λ then $E[X] = \frac{1}{\lambda}$

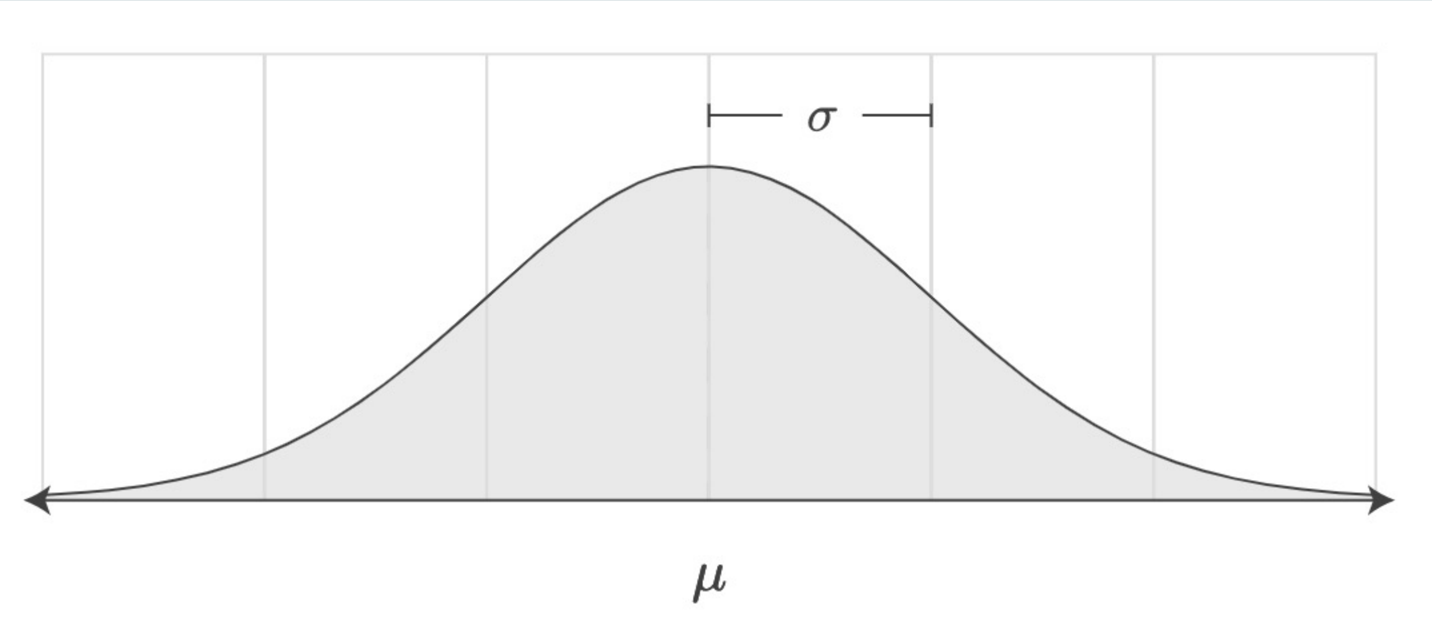
$$\boxed{E[X] = \int_{-\infty}^{\infty} xf(x)\,dx} \quad \Rightarrow \quad \boxed{E[X] = \int_0^{\infty} x\lambda e^{-\lambda x}\,dx = \frac{1}{\lambda}}$$

Notice the exponential distribution is the continuous analog of the geometric distribution.

The expectation of a normal distribution:

Let X be an $N(\mu, \sigma^2)$ distributed random variable, then: $E[X] = \mu$

$$\boxed{E[X] = \int_{-\infty}^{\infty} x f(x)\, dx} \Rightarrow \boxed{E[X] = \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = \mu}$$



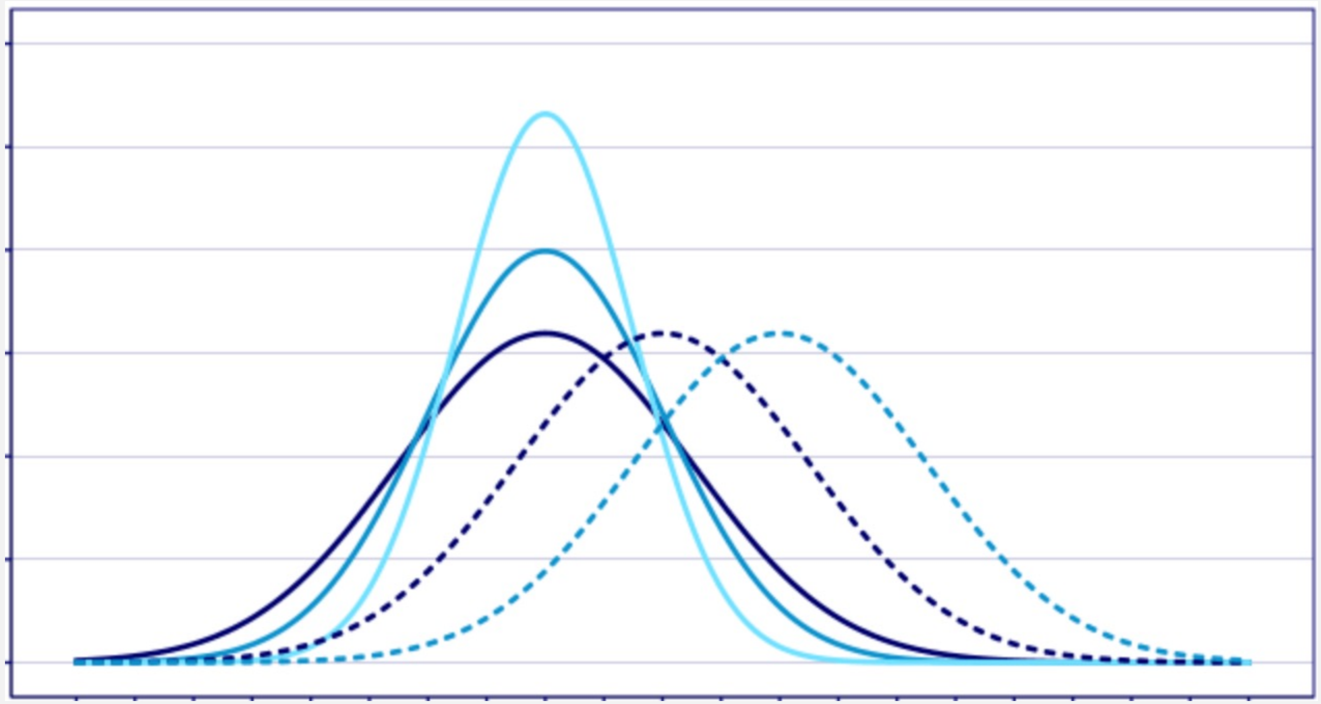
Suppose a random variable X has a given distribution.

If X is manipulated, then do the new values maintain the given distribution?

How is the expected value (mean) changed?

How is the variance changed?

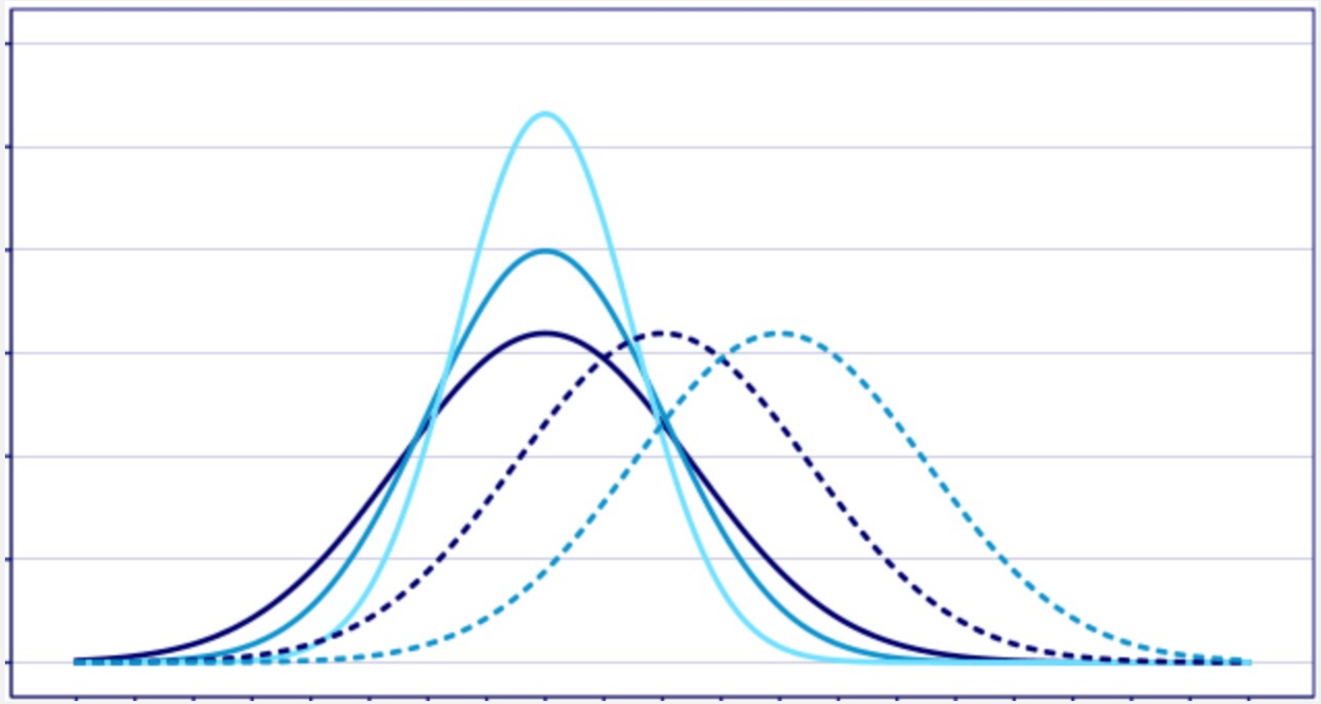
Does the new distribution become skewed?



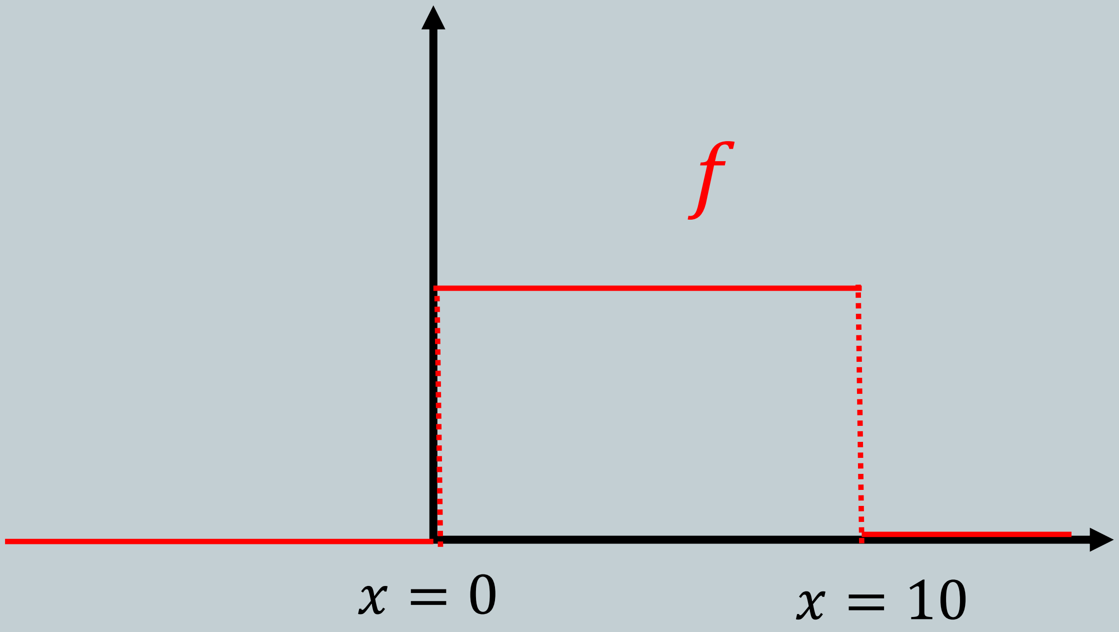
For instance, what happens to the distribution when you add n to each element?

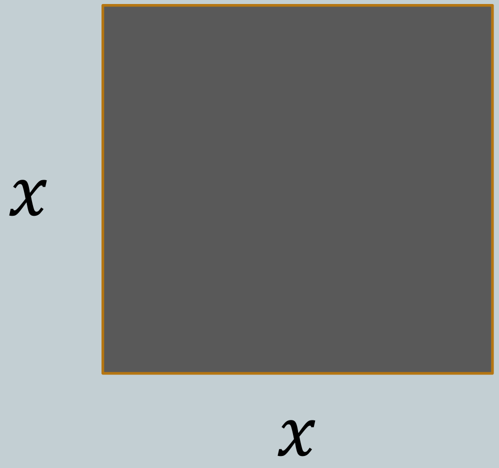
What if you multiplied each element by n ?

How do the expectation (mean) and variance change under addition or multiplication?



Suppose X is the side length of a square concrete foundation for a number of buildings on a building site. In an attempt to have an evenly disbursed variety of building sizes, X is uniformly distributed between 0 and 10 meters.





So X is uniformly distributed, but does that mean X^2 is also uniformly distributed?

Let $Y = X^2$ and let's find f_Y aka the distribution of building foundation area.

X is uniformly distributed between 1 and 10, therefore $f(x) = \frac{1}{10}$

$Y = X^2$ then has to be distributed between 0 and 100.
But distributed how? Also, uniformly?

We seek the distribution of area: f_Y

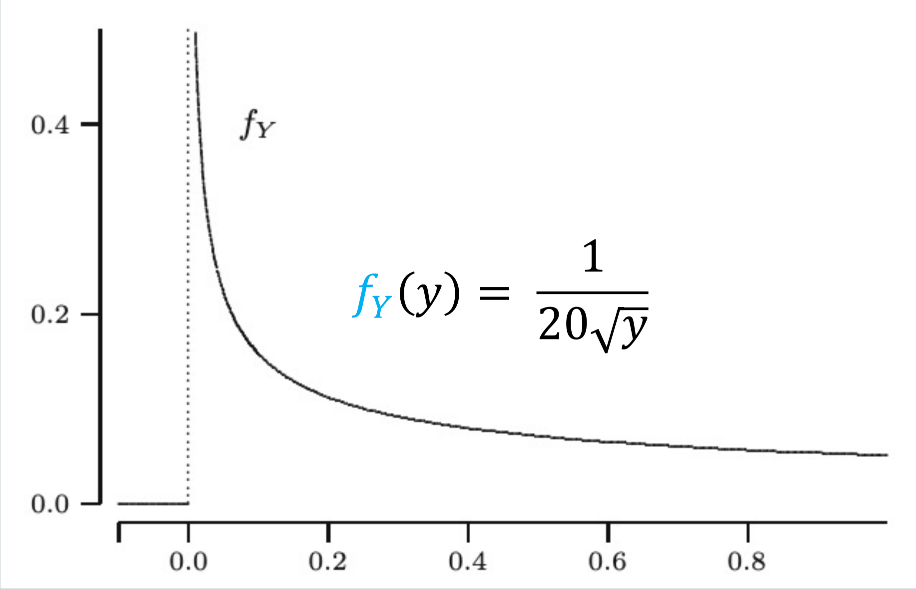
$$F_Y(a) = P(X^2 \leq a) = P(X \leq \sqrt{a}) = \frac{\sqrt{a}}{10}$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} \frac{\sqrt{y}}{10} = \frac{1}{20\sqrt{y}}$$

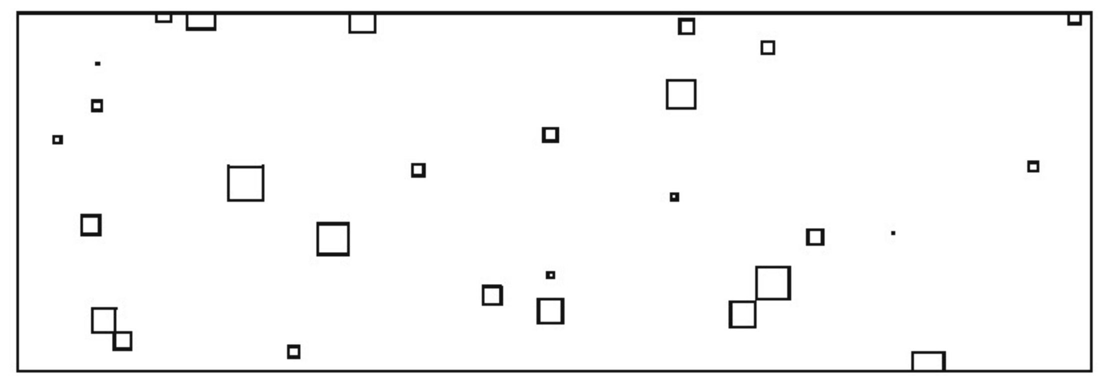
So… what?

The building areas are NOT uniformly distributed.

Actually, the buildings with small areas are heavily overrepresented as f_Y is asymptotic near 0.



Interestingly, this isn't obvious when looking at the overhead layout of the buildings selected with a Poisson process.



If you are a contractor making a bid on the concrete needed to pour foundations, then the amount of concrete needed is proportional to the area X^2 of a building. What do you expect for the area of the general building?

The expected area:

$$E[X^2] = E[Y] = \int_0^{100} y \cdot \frac{1}{20\sqrt{y}}\, dy = \int_0^{100} \frac{\sqrt{y}}{20}\, dy = \frac{1}{20}\frac{2}{3} y^{3/2} \Big|_0^{100} = 33\frac{1}{3}\ \ \mathrm{m}^2$$

Notice that the expected width and length of each foundation is 5,
(from the uniform distribution between 1 and 10)
but $5 \cdot 5 = 25\ \ m^2$ is not the correct expectation.

So, although you cannot find expected value of $\mathrm{E}[X^2]$ from $E[X] \cdot E[X]$, it is mathematically true that $\mathrm{E}[X^2]$ can be found from weighted average of those values:

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^{10} x^2 \cdot \frac{1}{10}\, dx = \frac{1}{30} x^3 \Big|_0^{10} = 33\frac{1}{3}\ \ \mathrm{m}^2$$

This leads us to the change-of-base formula:

Let X be a random variable.
Let $g: \mathbb{R} \to \mathbb{R}$ be a function.

If X is discrete, taking the values $a_1, a_2, \ldots$ then

$$E[g(X)] = \sum_i g(a_i)\, P(X = a_i)$$

If X is continuous, with probability density function f, then

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

These are used in practice when changing units.
Consider changing a random variable from Fahrenheit to Celsius.
What happens to the expectation?

Suppose X has a continuous distribution (such as Fahrenheit to Celcius would)
Apply the previous page formula to $g(x) = rx + s$ (r and s are $\mathbb{R}$).

$$E[g(X)] = \int_{-\infty}^{\infty} (rx + s) f(x)\, dx = r \int_{-\infty}^{\infty} x f(x)\, dx + s \int_{-\infty}^{\infty} f(x)\, dx = rE[X] + s$$

Let X have a $Ber(p)$ distribution. Compute $E[2^X]$.

ANS:

$$E[2^X] = \sum_i 2^{a_i} P(X = a_i)$$

$$= 2^0 \cdot P(X = 0) + 2^1 \cdot P(X = 1)$$

$$= 1 \cdot (1 - p) + 2 \cdot p$$

$$= 1 - p + 2p$$

$$= 1 + p$$

Alternatively:

$Y = 2^X$ has a distribution given by $P(Y = 1) = 1 - p$ and $P(Y = 2) = p$.

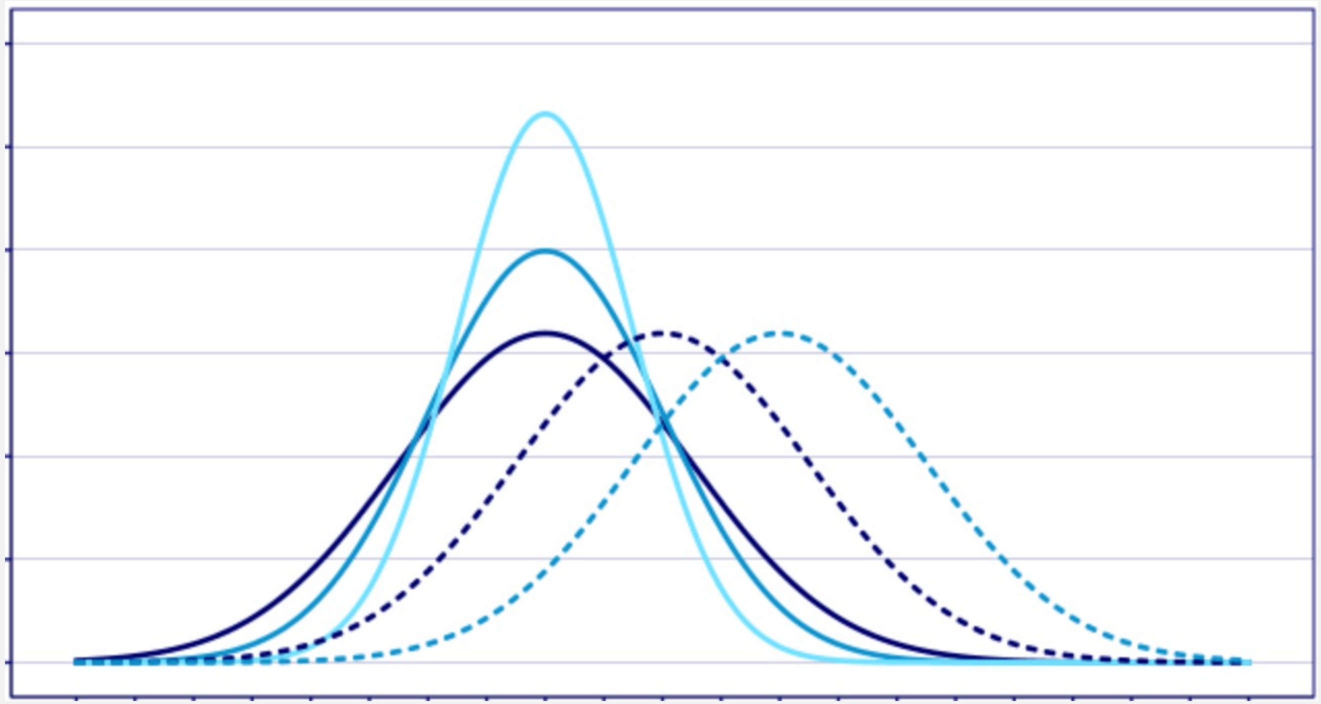
Therefore:

$$E[2^X] = E[Y] = 1 \cdot P(Y = 1) + 2 \cdot P(Y = 2) = 1 \cdot (1 - p) + 2 \cdot p = 1 + p$$

We spoke of summarizing a distribution with one number.

- Our first/best answer was the expected value or mean.

- Another descriptor is the variance



Suppose you have the opportunity to make an investment in one of two ventures. Which would you pick? Well, what is the expected return?

Investment 1: Expected return is \$500.

Investment 2: Expected return is \$500.

They sound the same until you get some extra information:
Investment 1 has a 50% chance of returning \$450 and a 50% chance of returning \$550.
Investment 2 has a 50% chance of returning \$0 and a 50% chance of returning \$1000.

Although both investments have the same expected value of \$500:

$\frac{450+550}{2} = 500$ and $\frac{0+1000}{2} = 500$, they have very different variances.

The spread around the mean of a random variable is of great importance.
This spread is measured by the expected squared deviation from the mean.

This spread is measured by the expected squared deviation from the mean

Definition:

The variance of a random variable X is the number: $\mathrm{Var}(X) = E[(X - E[X])^2]$.

Standard Deviation: $\sqrt{\mathrm{Var}(X)}$ is useful as it has the same dimension as $E[X]$.

If we apply the change of base formula to $\mathrm{Var}(X)$, then we arrive at a somewhat faster alternative expression for variance.

Suppose X is a continuous random variable with probability density function f, then:

$$\begin{aligned}
\mathrm{Var}(X) &= E[(X - E[X])^2] \\
&= \int_{-\infty}^{\infty} (x - E[X])^2 f(x)dx \\
&= \int_{-\infty}^{\infty} (x^2 - 2xE[X] + (E[X])^2) f(x)dx \\
&= \int_{-\infty}^{\infty} x^2 f(x)dx - 2E[X] \int_{-\infty}^{\infty} xf(x)\, dx + (E[X])^2 \int_{-\infty}^{\infty} f(x)\, dx \\
&= E[X^2] - 2(E[X])^2 + (E[X])^2 \\
&= E[X^2] - (E[X])^2
\end{aligned}$$

This is useful in that it is often not practical to compute $\mathrm{Var}(X)$ directly from the definition.

Recall the drilling example:

Hours	Probability
2	.10
3	.70
4	.20

$$\mathrm{Var}(X) = E[(X - E[X])^2]$$

$$\begin{aligned} \mathrm{Var}(\mathrm{X}) &= E[(X - 3.1)^2] \\ &= 0.1 \cdot (2 - 3.1)^2 + 0.7 \cdot (3 - 3.1)^2 + 0.2 \cdot (4 - 3.1)^2 \\ &= \mathbf{0.29} \end{aligned}$$

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2$$

$$\begin{aligned} \mathrm{Var}(\mathrm{X}) &= E[X^2] - (3.1)^2 \\ &= 0.1 \cdot 2^2 + 0.7 \cdot 3^2 + 0.2 \cdot 4^2 - 9.61 \\ &= \mathbf{0.29} \end{aligned}$$

Recall the investment example:

Call the return from the first investment, Y_1

Call the return from the second investment, Y_2

What are the variance and standard deviation of Y_1 and Y_2 ?

ANS:

$$\mathrm{Var}(Y_1) = \frac{1}{2}(450 - 500)^2 + \frac{1}{2}(550 - 500)^2 = 50^2 = 2500$$

Standard Deviation of Y_1 is 50.

$$\mathrm{Var}(Y_2) = \frac{1}{2}(0 - 500)^2 + \frac{1}{2}(1000 - 500)^2 = 500^2 = 250{,}000$$

Standard Deviation of Y_2 is 500.

$$Var(rX) = E[(rX - E[rX])^2]$$

$$= E[(rX - rE[X])^2]$$

$$= E[r^2(X - E[X])^2]$$

$$= r^2 E[(X - E[X])^2]$$

$$= r^2 Var(X)$$

$$Var(X+s) = E\left[\left((X+s) - E[X+s]\right)^2\right]$$

$$= \left[\left((X+s) - E[X] - s\right)^2\right]$$

$$= E[(X - E[X])^2]$$

$$= Var(X)$$

Therefore:

$$Var(rX+s) = Var(rX) = r^2 Var(X)$$

Recall our discussion on change of units (Fahrenheit to Celcius).
It was noted that $E[rX+s] = rE[X]+s$.

We can arrive at the corresponding rule for the variance under change of units.

Expectation and variance under change of units

For any random variable X and any real numbers r and s,

$$E[rX+s] = rE[X]+s \quad \text{and} \quad \mathrm{Var}(rX+s) = r^2\mathrm{Var}(X)$$

Notice that variance is not affected by s.
What does that mean?

In Summary:

$E[X] = \sum_i a_i P(X = a_i) = \sum_i a_i\, p(a_i)$ and $E[X] = \int_{-\infty}^{\infty} x f(x)\, dx$

$X \sim Ber[p] \rightarrow E(X) = p \qquad Var(X) = p(1-p)$

$X \sim Bin(n,p) \rightarrow E[X] = np \qquad Var(X) = np(1-p)$

$X \sim U[\alpha,\beta] \rightarrow E[X] = \frac{1}{2}(\alpha+\beta) \qquad Var(X) = \frac{1}{12}(\beta-\alpha)^2$

$X \sim N(\mu,\sigma^2) \rightarrow E[X] = \mu \qquad \mathrm{Var}(X) = \sigma^2$

Why? Try calculating them by definition to see if you get the same results listed.

If $X \sim U[\alpha, \beta]$, then $E[X] = \frac{1}{2}(\alpha + \beta)$ and $Var(X) = \frac{1}{12}(\beta - \alpha)^2$

(Recall: $E[X] = \sum_i a_i P(X = a_i)$)

If X is discrete, taking the values $a_1, a_2, \ldots$ then

$$E[g(X)] = \sum_i g(a_i) P(X = a_i)$$

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

If X is continuous, with probability density function f, then

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

Let X have a $Ber(p)$ distribution. Compute $E[2^X]$.

ANS:

$$E[2^X] = \sum_i 2^{a_i} P(X = a_i)$$
$$= 2^0 \cdot P(X = 0) + 2^1 \cdot P(X = 1)$$
$$= 1 \cdot (1 - p) + 2 \cdot p$$
$$= 1 - p + 2p$$
$$= 1 + p$$

Alternatively:
$Y = 2^X$ has a distribution given by $P(Y = 1) = 1 - p$ and $P(Y = 2) = p$.
Therefore:

$$E[2^X] = E[Y] = 1 \cdot P(Y = 1) + 2 \cdot P(Y = 2) = 1 \cdot (1 - p) + 2 \cdot p = 1 + p$$

Expectation and variance under change of units

For any random variable X and any real numbers r and s,

$$E[rX+s] = rE[X]+s \qquad \text{and} \qquad \mathrm{Var}(rX+s) = r^2\mathrm{Var}(X)$$

Alternat expression for variance.

Suppose X is a continuous random variable with probability density function f, then:

$$\mathrm{Var}(X) \quad = \quad E[(X - E[X])^2] \qquad = \quad E[X^2] - \ (E[X])^2$$

Next time: More on the normal distribution and the central limit theorem