Log-Normal distribution

Consider an example of number of days of hospitalisation

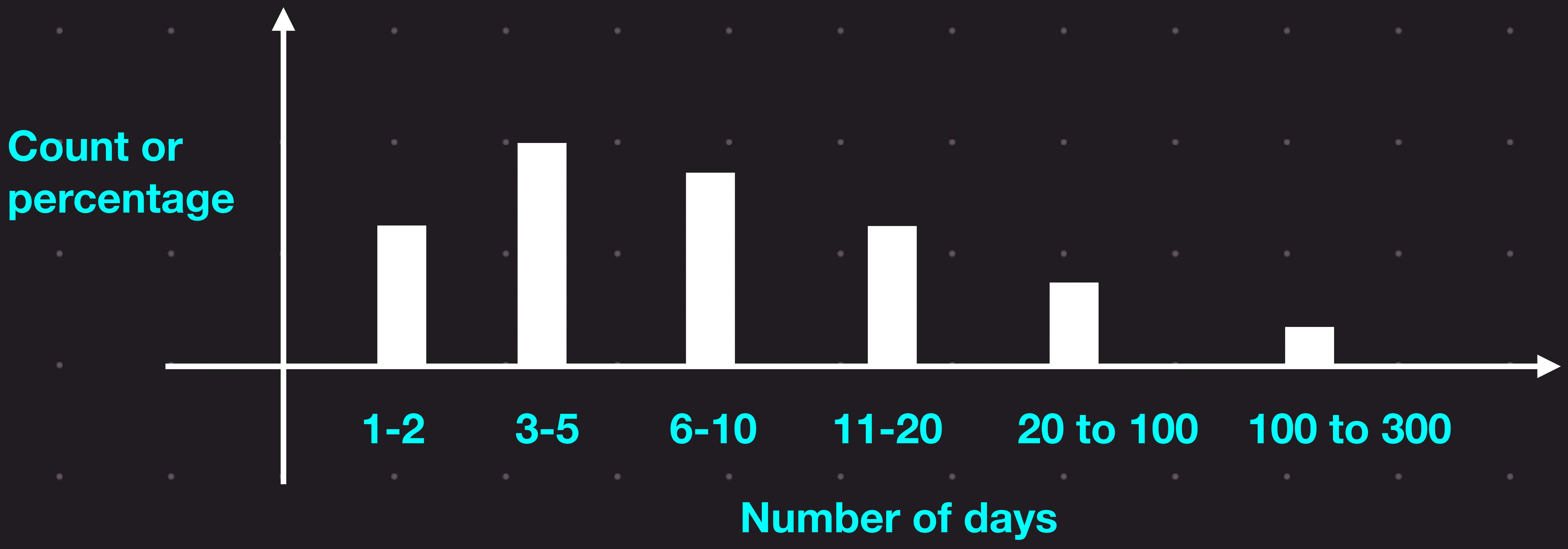
Most people may be hospitalised for 3 to 5 days

Quite often, just 1 or 2 days

But, there can be a few extreme cases in 20 to 30 days

Extremely rare cases of 300 days (coma etc.)

Log-Normal distribution



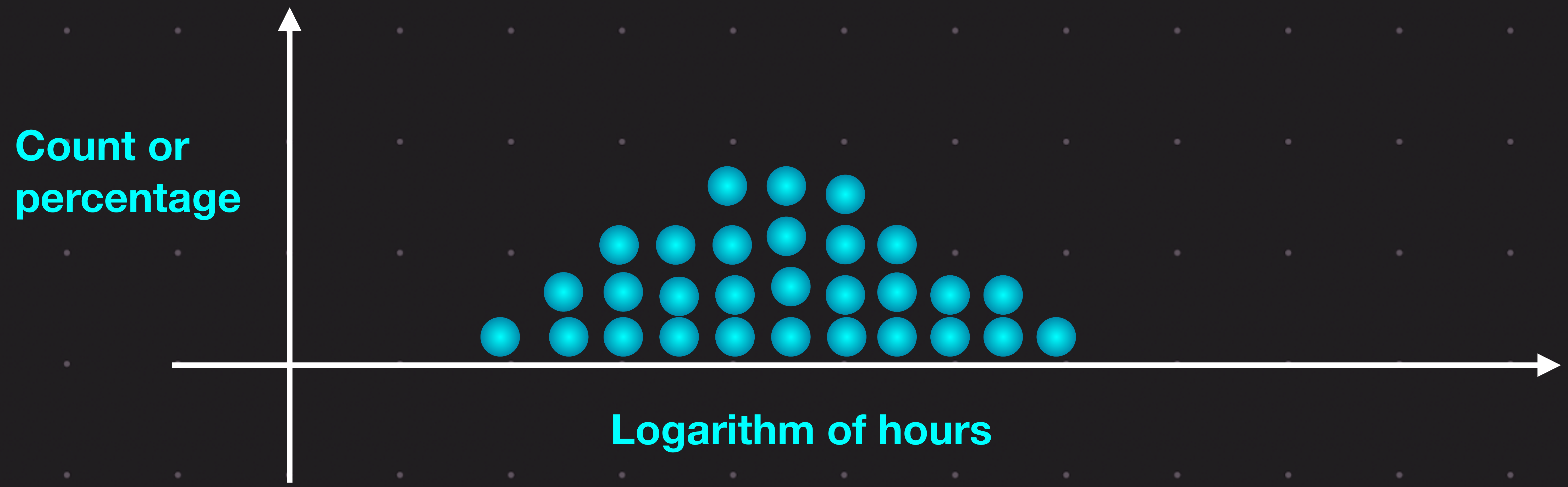
Number of days is discrete

What if we had exact number of hours?

This would be continuous

Log-Normal distribution

Log-Normal distribution



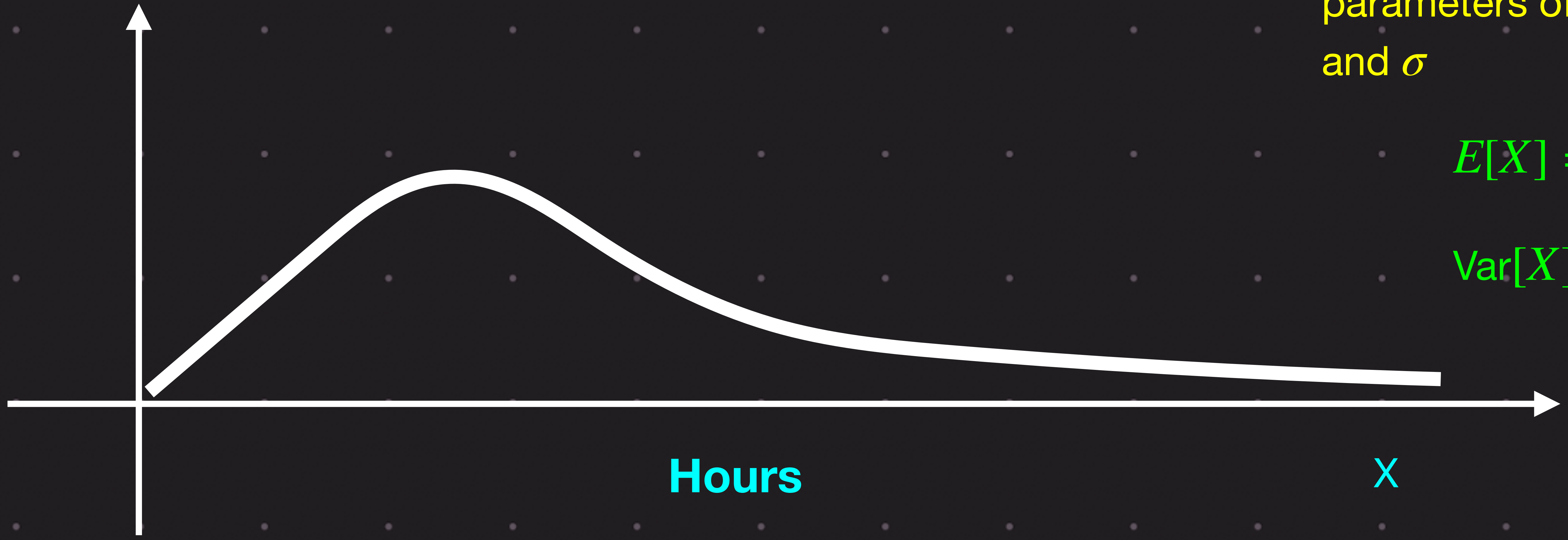
Log-Normal distribution

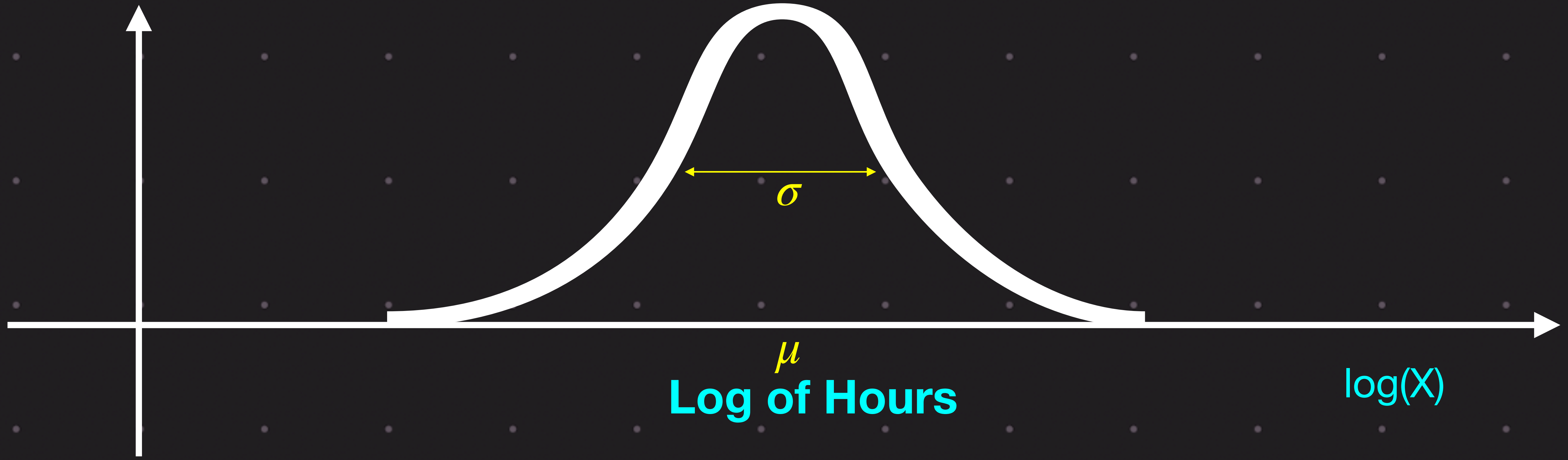
For the log normal, we use the parameters of the Gaussian: μ and σ

$$E[X] = e^{\mu+\sigma^2/2}$$

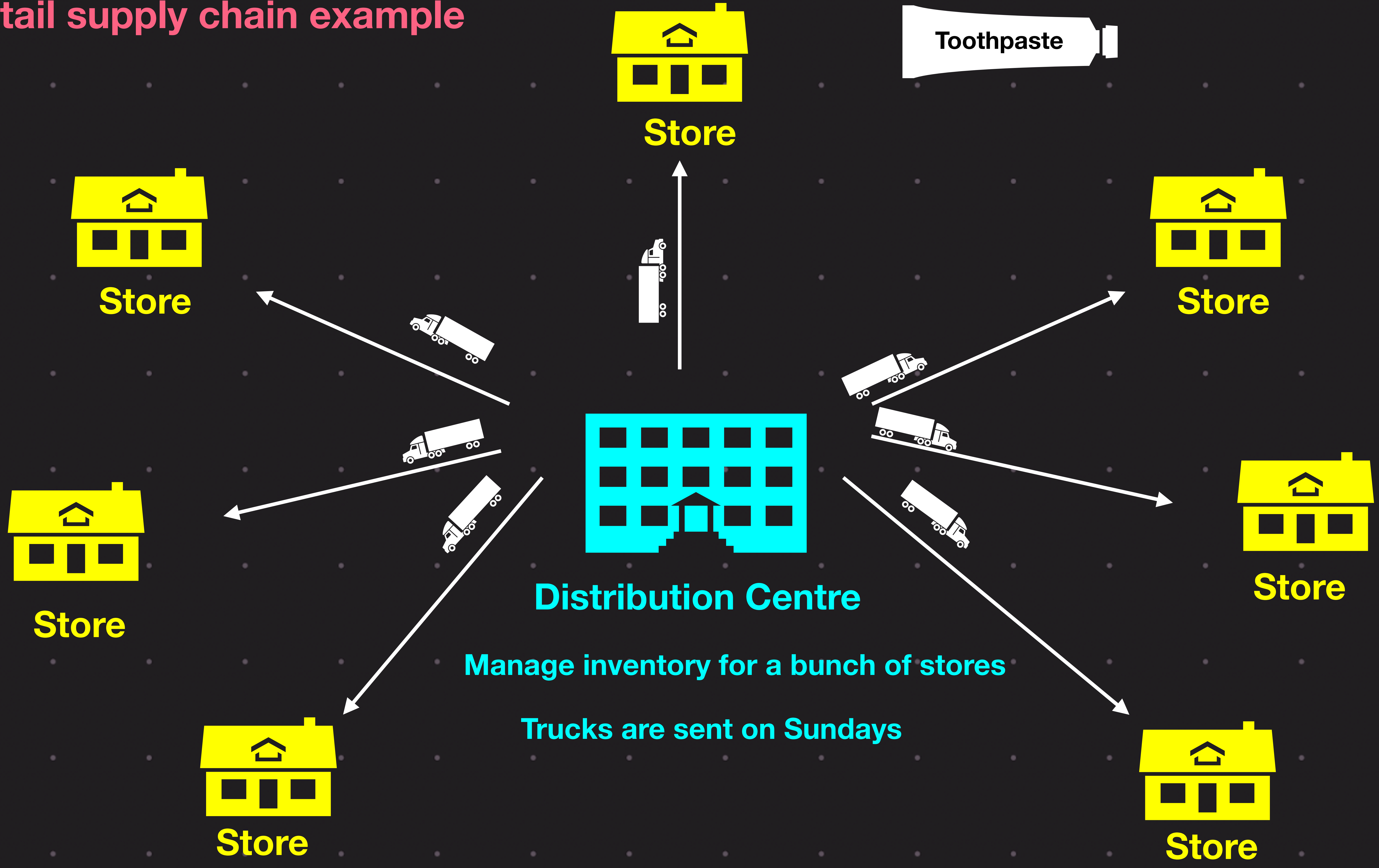
$$\text{Var}[X] = (e^{\sigma^2} - 1)\left(e^{2\mu+\sigma^2}\right)$$

Always confirm these equations from some source (eg: wiki)





Retail supply chain example



Manage inventory for a bunch of stores

Trucks are sent on Sundays

Airline Overbooking

Five percent of the people making reservations on a flight will not show up.
Suppose the airline sells 52 tickets for a flight that can hold only 50 passengers.
What is the probability that there will be a seat available for every passenger who shows up?

Probability of showing up: $p = 0.95$

Let X denote the number of people who show up

What are we asked to find? $P[X \leq 50]$

Can we convert this to asking this question:
A coin is tossed 52 times, the probability of heads is 0.95.
What is the probability of 50 or lesser heads?

$P[X \leq 50] =$ `binom.cdf(k=50, n=52, p=0.95)` $= 0.74$

Pooled Blood test

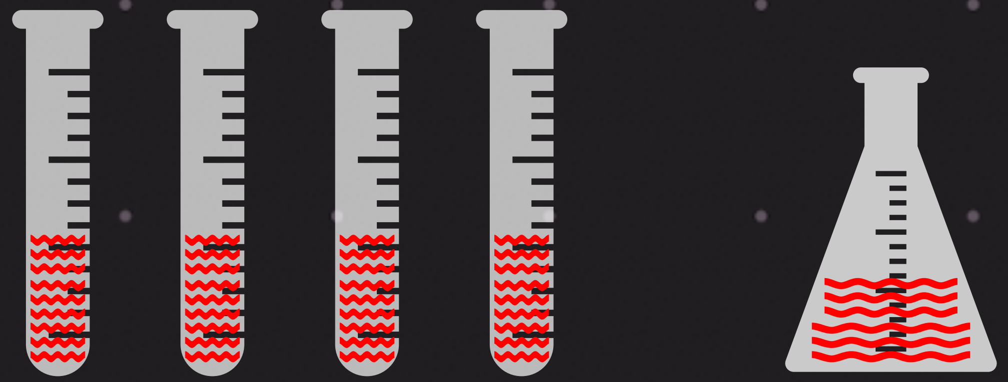
A blood bank tests pooled samples of 4 people at a time.

If clean, the bank stores all four.

If unacceptable, then all 4 samples are tested individually.

The probability of any sample being unacceptable is 0.1.

Find the expected number of tests needed



Probability of individual begin unacceptable: $p = 0.1$

Let X denote the number of tests needed

What is the question asking? $E[X]$

What are the values that X can take? 1 or 5

$P[X=1] =$ `binom.pmf(k=0, n=4, p=0.1)` $= 0.6561$

$$P[X=1] = {}^{4}C_{0}(0.1)^{0}(0.9)^{4} = (0.9)^{4} = 0.6561$$

$$P[X=5] = 1 - 0.6561 = 0.3439$$

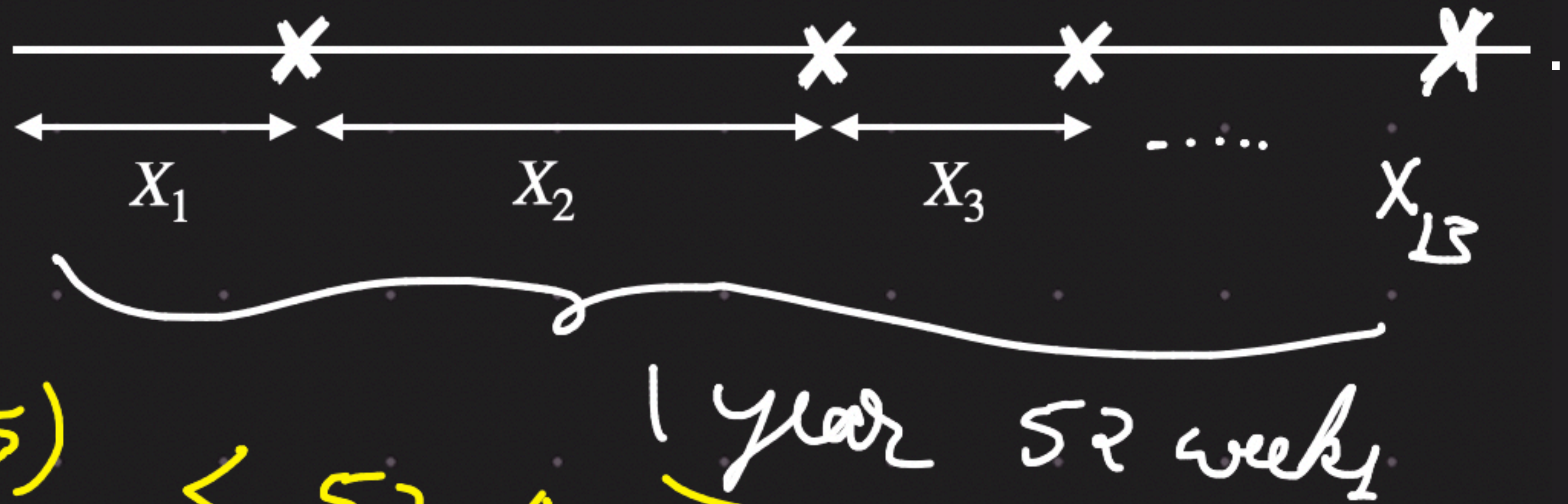
$$\begin{aligned} E[X] &= 1 * P[X=1] + 5 * P[X=5] \\ &= 1 * 0.6561 + 5 * 0.3439 \\ &= 2.37 \end{aligned}$$

Machine Learning production failure

An ML algorithm fails in production with an average of 5 weeks and std dev 1.5 weeks

Assume these failures are immediately fixed by the on-call support team

Find the probability of 13 or more failures in a year



$$Y = X_1 + X_2 + \cdots + X_{13} \qquad (1)$$

$$P[Y < 52] = P\left[\frac{Y-(13)(5)}{\sqrt{13}\,(1.5)} < \frac{52-(13)(5)}{\sqrt{13}\,(1.5)}\right]$$

$$= P\left[Z < (\quad)\right]$$

$\mu = 5$, $\sigma = 1.5$

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_{13}}{13} \quad (1)$$ → " sample mean "

$$E[\bar{X}] = 5 \quad (2)$$

$$\text{Std dev}_{\bar{X}} = \frac{\sigma}{\sqrt{13}} = \frac{1.5}{\sqrt{13}} \quad (3)$$

$$Y = 13\,\bar{X} \quad (4)$$

$$\sigma_Y = 13\,\sigma_{\bar{X}} \quad (5)$$

$$Y = X_1 + X_2 + \cdots + X_{13} \quad (4)$$

$$E[Y] = (13)\,\mu = \underline{13(5)} \quad (6)$$

$$\sigma_Y = \underline{13\left(\frac{1.5}{\sqrt{13}}\right)} \quad (7)$$

Simulate a fair coin from a biased coin

There is a coin that lands heads 70% of the times

How can we use this coin so that it lands heads 50% of the times?

Let us toss the biased coin twice

Sample space S = {HH, HT, TH, TT}

$$P[HH] = 0.7 * 0.7$$
$$P[HT] = 0.7 * 0.3$$
$$P[TH] = 0.3 * 0.7$$
$$P[TT] = 0.3 * 0.3$$

These two have same probability

Biased	Fair
$HH \rightarrow$	Ignore
$TT \rightarrow$	Ignore
$HT \rightarrow$	H
$TH \rightarrow$	T