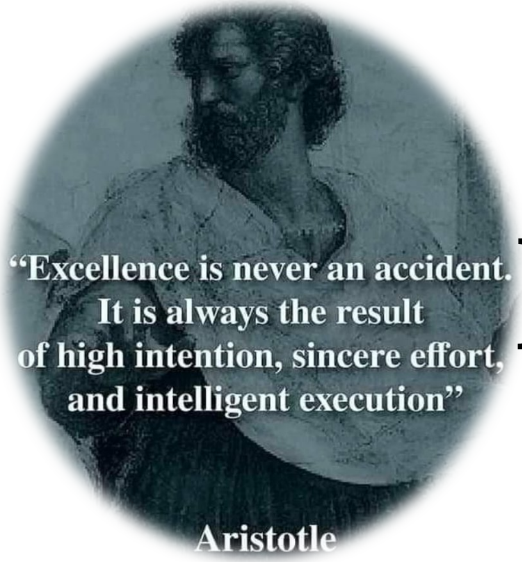
combra.cs.rutgers.edu



206
Discrete Structures II

Konstantinos P. Michmizos

Computational Brain Lab

Computer Science | Rutgers University | NJ, USA

Lecture 24 | Random Variables and Expectations | Monday December 11th 2023

CTAAR survey

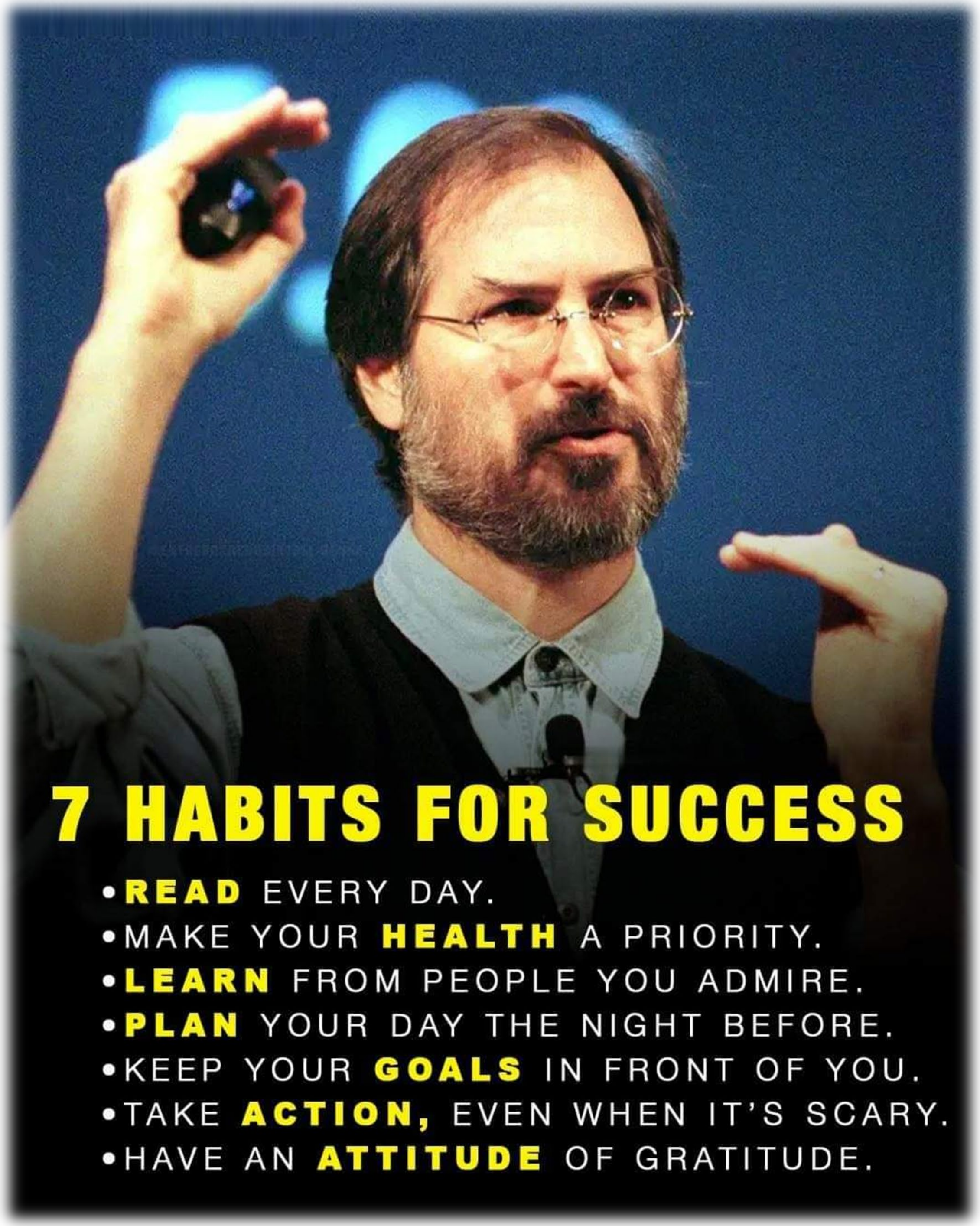
https://sirs.rutgers.edu/blue

CTAAR Survey

Section 1

Section 2

Section 3


7 HABITS FOR SUCCESS
•READ EVERY DAY.
•MAKE YOUR HEALTH A PRIORITY.
•LEARN FROM PEOPLE YOU ADMIRE.
•PLAN YOUR DAY THE NIGHT BEFORE.
•KEEP YOUR GOALS IN FRONT OF YOU.
•TAKE ACTION, EVEN WHEN IT'S SCARY.
•HAVE AN ATTITUDE OF GRATITUDE.

This is the...Last Week of Lectures

Part 1: Counting

- Proofs/induction
- Sum Rule
- Partition Method
- Difference Method
- Bijection Method
- Product Rule
- Generalized product rule
- Pigeonhole Principle
- Inclusion/Exclusion
- Combinatorial proofs, binomial coefficients

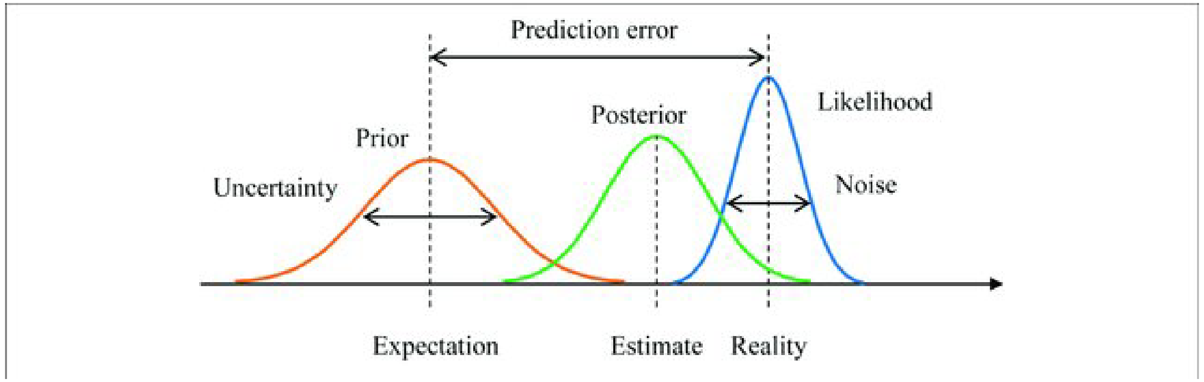
Part 2: Probability

- Sample space
- Axioms of probability
- Conditional probability
- Independence
- Bayes rule
- Random Variables
- Expectation and Linearity
- Variance

Bayesian Inference*

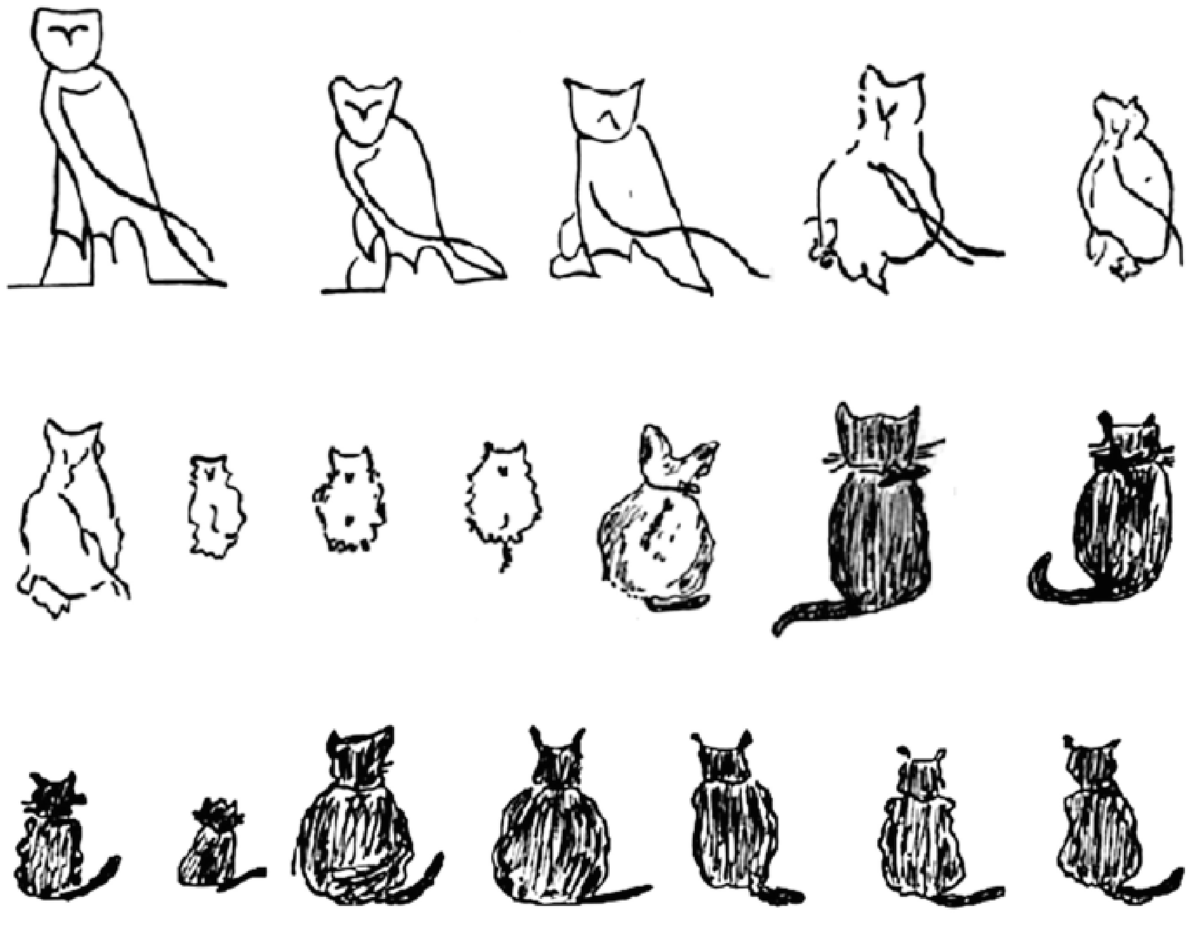
$$\overset{\text{Posterior}}{P(A|B)} = \frac{\overset{\text{Likelihood}}{P(B|A)} * \overset{\text{Prior}}{P(A)}}{\underset{\text{Evidence}}{P(B)}}$$

* *Inference = Educated guessing*



- Bayesian inference with a prior distribution, a posterior distribution, and a likelihood function.
- The prediction error is the difference between the prior expectation and the peak of the likelihood function (i.e., reality).
- Uncertainty is the variance of the prior. Noise is the variance of the likelihood function.

But do we really understand Bayes Rule?

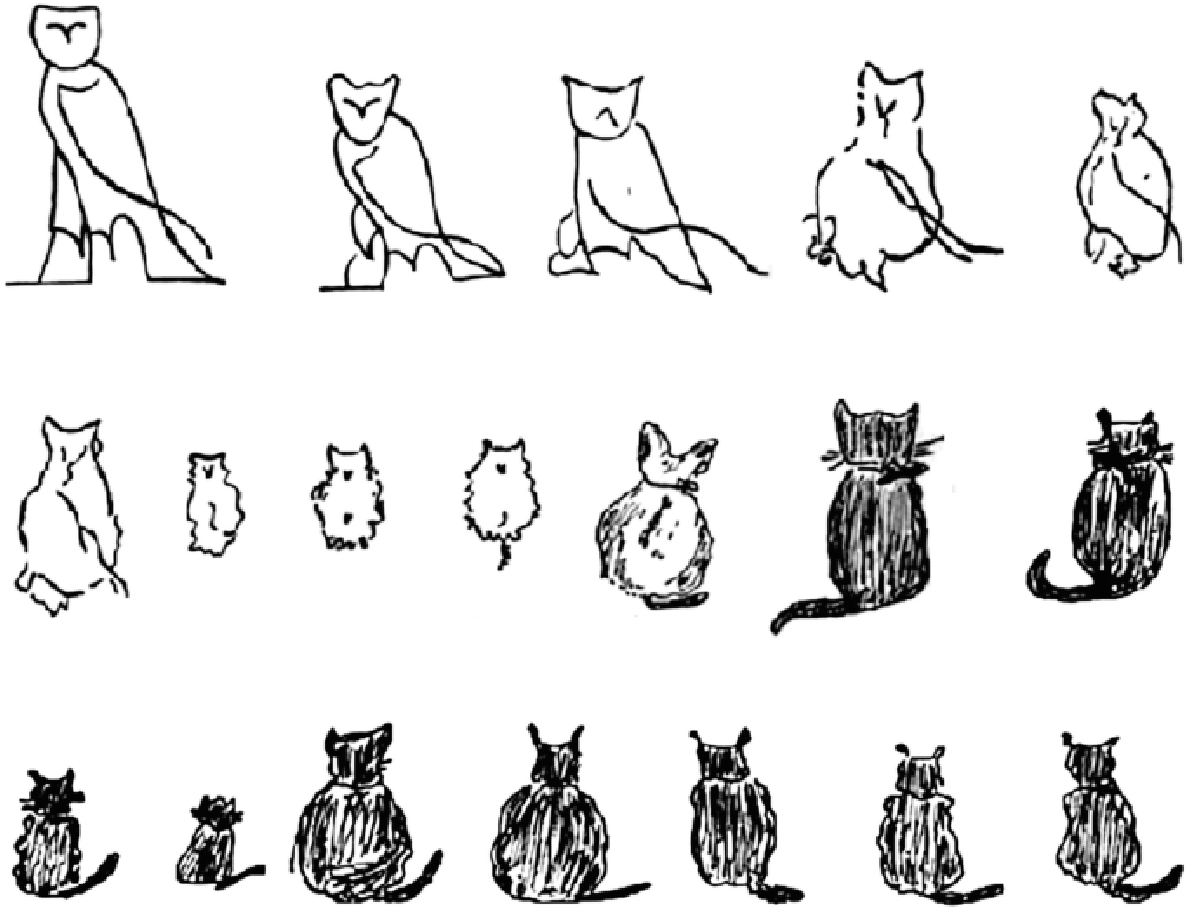


Capturing the *assimilation* of unfamiliar material to conventional cultural patterns.

Bartlet et al. 1932

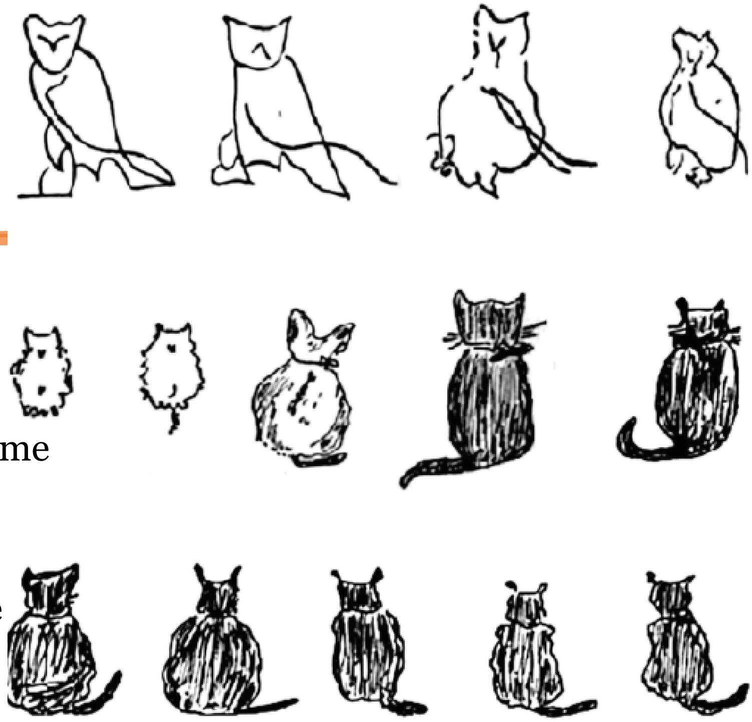
The method begins with showing a participant a story or image from foreign culture, which is thus unfamiliar to them.

After fifteen minutes has elapsed, a participant is asked to reproduce the material from memory. Their reproduction is then shown to a second person, who does the same, and so on.



Serial Reproduction: What happened to the ancient Egyptian hieroglyph of an owl (top left corner) when people living in England in the 1910s serially reproduced it

Bartlet et al. 1932



Early reproductions: The design became more oval shaped and its inner features became less clearly connected.

Eighth reproduction: A tail appears, suggesting that the participant interpreted the design to represent some kind of furry animal.

Ninth reproduction: Much more distinctly cat like: The ears grown in size and the inner markings become shading and a collar.

Next reproductions: The design is clearly of a black cat, where even whiskers are added.

Last reproductions: Then very little change in reproductions to the end of the series, besides the moving of the tail.

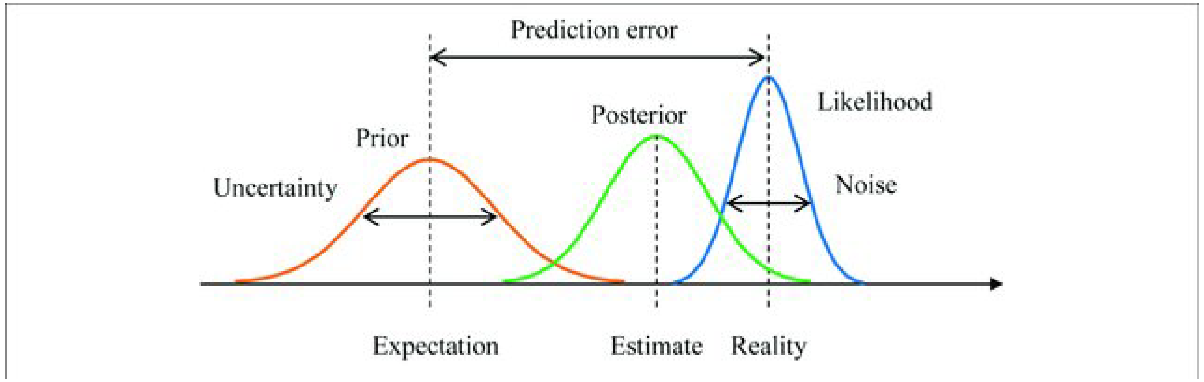
At this point, the design has become a conventional representation in the culture it is circulating in. *In contrast to the original owl figure, people recognize the cat immediately and can reproduce it rapidly*.

The foreignness of the evidence leads to greater changes as it is worked into the receiving group's cultural patterns.

Bayesian Inference*

$$\overset{\text{Posterior}}{\Downarrow}\; P(A|B) = \frac{\overset{\text{Likelihood}}{\Downarrow}\;P(B|A) * \overset{\text{Prior}}{\Downarrow}\;P(A)}{\underset{\text{Evidence}}{\Uparrow}\;P(B)}$$

** Inference = Educated guessing*



- Bayesian inference with a prior distribution, a posterior distribution, and a likelihood function.
- The prediction error is the difference between the prior expectation and the peak of the likelihood function (i.e., reality).
- Uncertainty is the variance of the prior. Noise is the variance of the likelihood function.

Today...

- We will understand the notion of expectation in random variables

$$E[X] = \sum_{e \in \Omega} X(e)\, P(e)$$

$$E[X] = \sum_{k} k \cdot P(X=k)$$

k ranges over all possible values that X can take.

Why Random Variables?

Compare this…

- A fair coin is tossed 100 times. What is the probability that we get exactly 50 heads?
- If we roll a white die and a black die (both fair), what is the probability that the sum is 7?

Why Random Variables?

...against this (meta-questions)

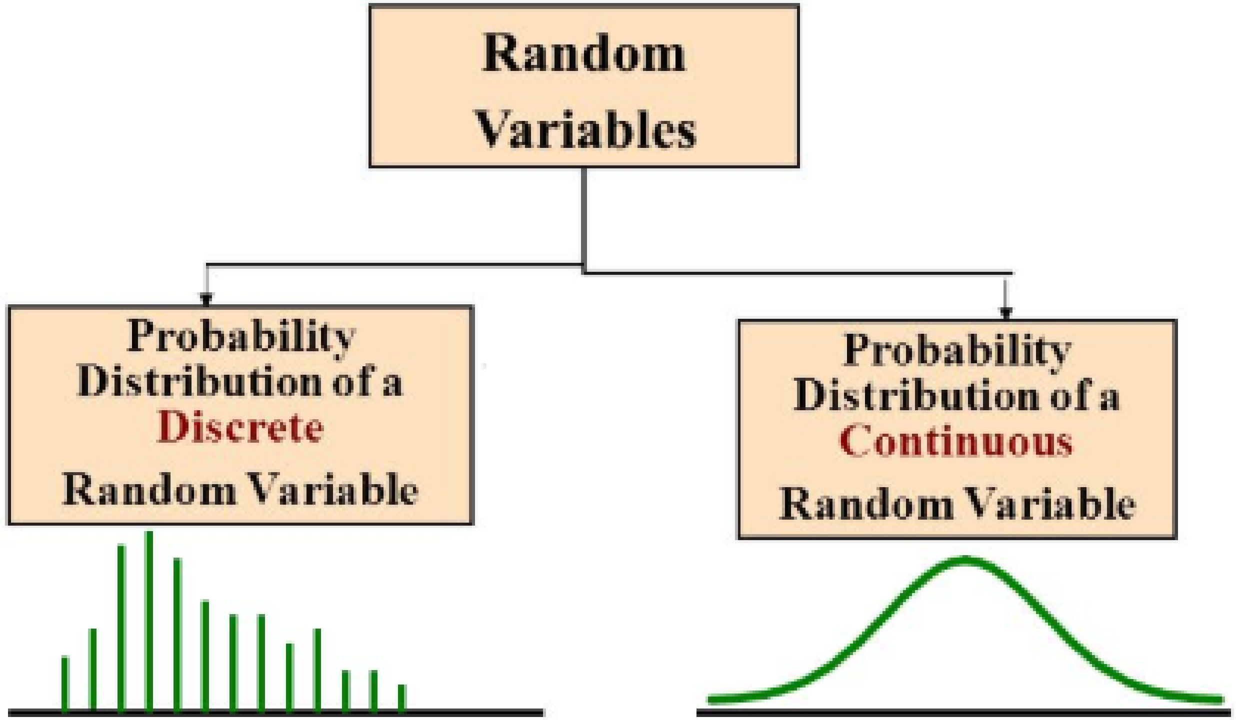
- A fair coin is tossed 100 times. What is the average number of heads seen?
- If we roll a white die and a black die (both fair), what is the average value of the sum?

This is different than asking for P(A) for some event A.

Meta-ambiguity

What is our Expectation for a given random variable?

Random Variables



What is not a random variable?

- It is not random
- It is not a variable
- So, what is it?
  - It is a function
- A Function of what?
  - A function of the sample set of an experiment
    - It associates each **outcome** of an experiment **with a real number**

Example of a Random ~~Variable~~ Function

EXAMPLE: 2 coins Toss $S = \{HH, HT, TH, TT\}$

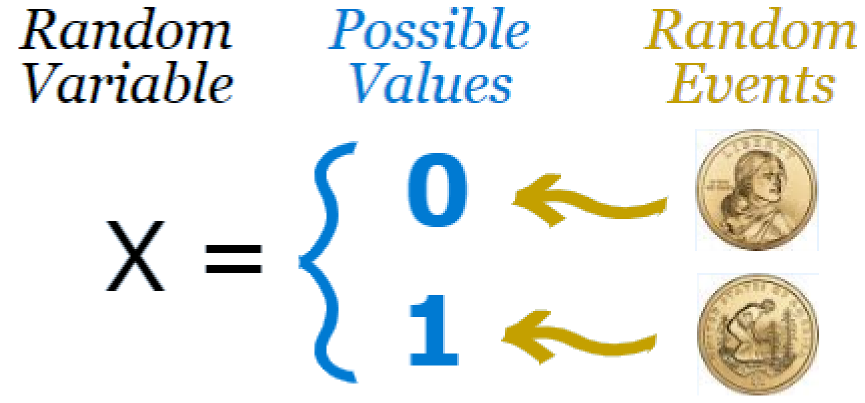
Name of function/Random Variable : X

Define Function: X : {Number of Heads}

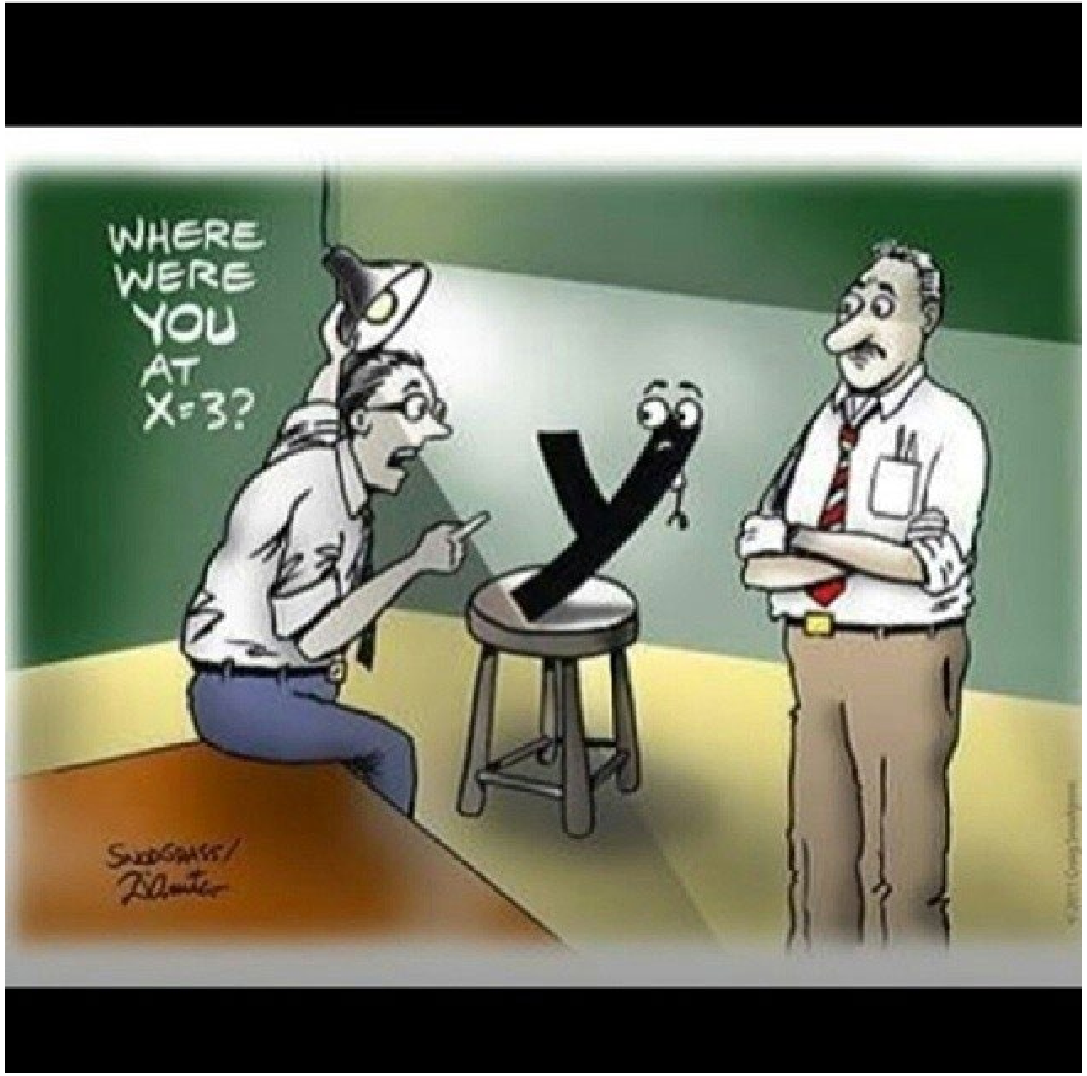
S	HH	HT	TH	TT
X	2	1	1	0

So... What is a random variable?

a random variable is
the **result**
of a **chance event**
that we measure or
count



X = {0, 1}



What is not a random variable?

- A Random Variable is Not Like an Algebra Variable

- In Algebra a variable, like **x,** is an unknown value:

e.g., **x** + 2 = 6 → x=4

- A Random Variable has a whole **set of values** and it could take on **any** of those values, randomly.

e.g., **X = {0, 1, 2, 3}**

X could be 0, 1, 2, or 3 *randomly*. And they might each *have a different probability*.

How do we annotate a random variable?

Capital Letters

- We use a capital letter, like **X** or **Y**, to avoid confusion with the Algebra type of variable.

A Random Variable’s Sample Space

- A Random Variable's set of possible values.

Example: Throw a die once

- Random Variable **X** = "The score shown on the top face".
- **X** could be 1, 2, 3, 4, 5 or 6
- So the Sample Space is {1, 2, 3, 4, 5, 6}

Probability of a value

Throw a die once

- Sample Space: $X = \{1, 2, 3, 4, 5, 6\}$

All values are equally likely, so the probability of any one is 1/6

- $P(X = 1) = 1/6$
- $P(X = 2) = 1/6$
- $P(X = 3) = 1/6$
- $P(X = 4) = 1/6$
- $P(X = 5) = 1/6$
- $P(X = 6) = 1/6$

A better example: Seif's ice cream stand

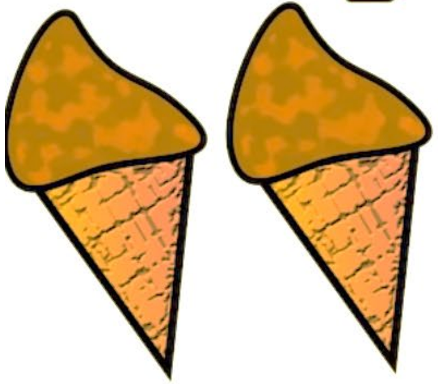
Seif sells ice cream in cones.

He wants to have a better idea of what is going on in his business:

- How many ice creams are typically bought in each transaction?
- How many single-ice cream customers?
- How many customers buy more than 3 cones?
- How many cones to stock for weekends?

etc.

Seif's ice cream stand: From data to R.V.



Number of ice creams	Customers
1	225
2	170
3	55
4	20
5	20
6	10

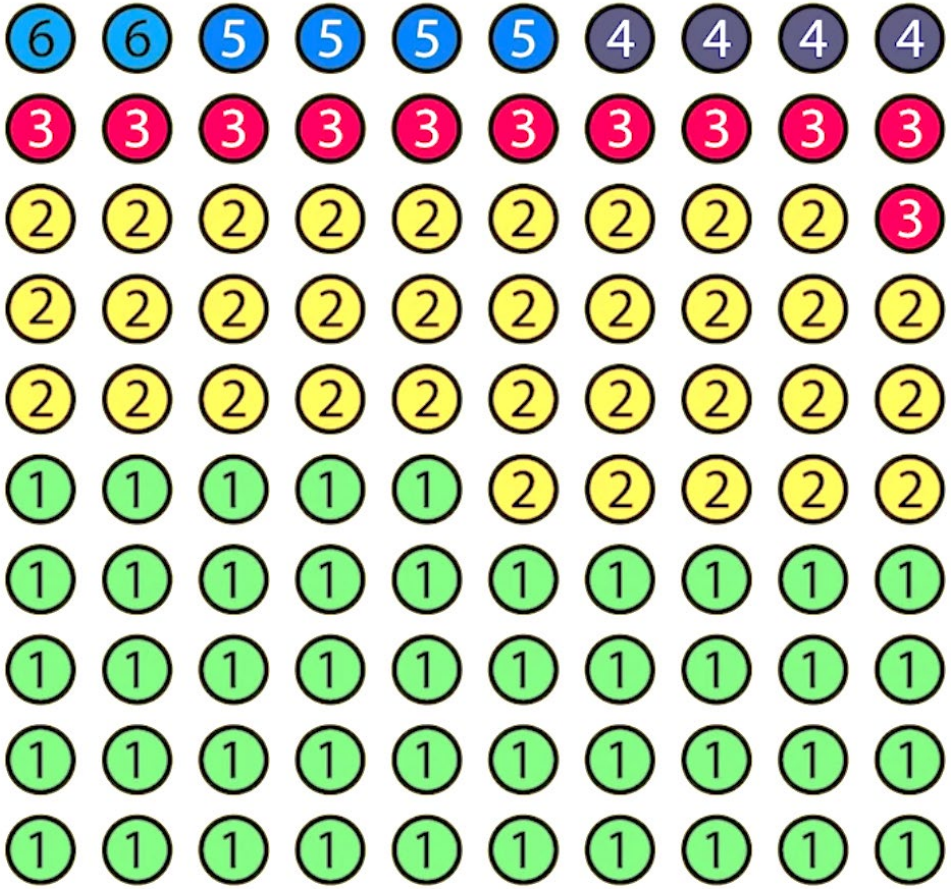
Seif's ice cream stand: From data to R.V.

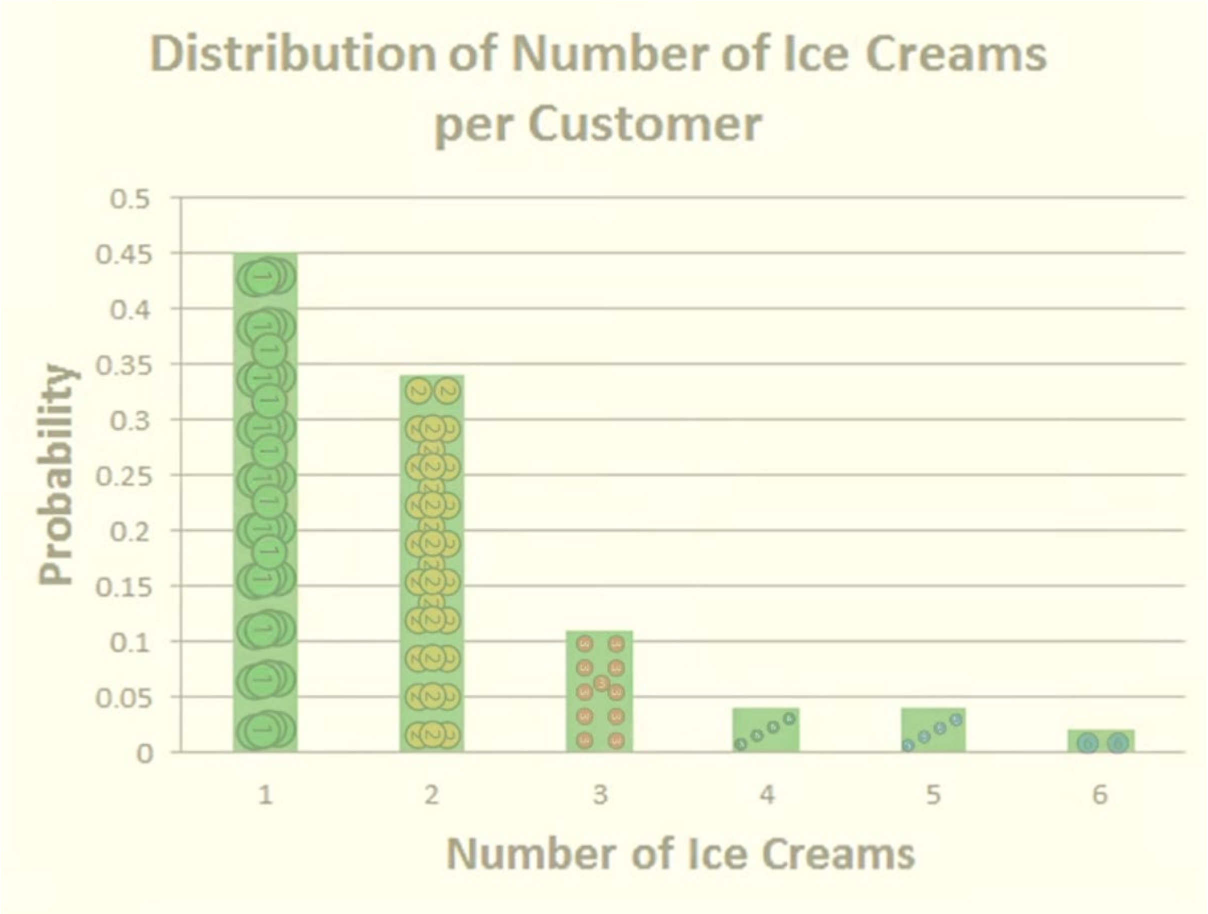
Random Variable

$X =$ number of ice creams a customer orders

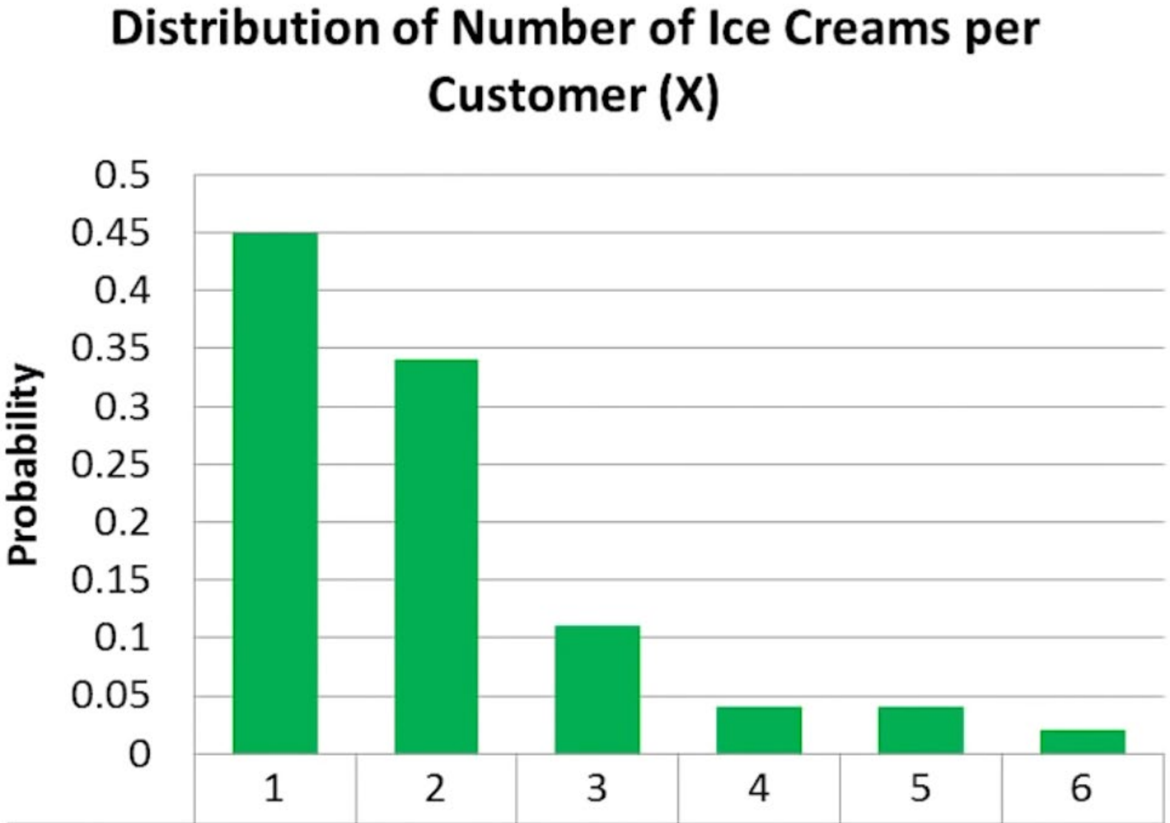
Number of ice creams (x)	Customers	P(X= x)
1	225	0.45
2	170	0.34
3	55	0.11
4	20	0.04
5	20	0.04
6	10	0.02
Total	500	1

Seif's ice cream stand: From data to R.V.





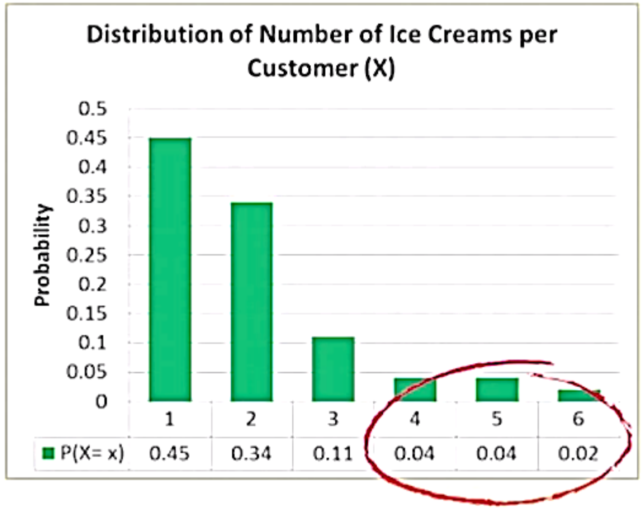
Asking Questions to a R.V.



	1	2	3	4	5	6
P(X= x)	0.45	0.34	0.11	0.04	0.04	0.02

Asking Questions to a R.V.

$$P(X>3)$$

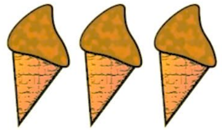


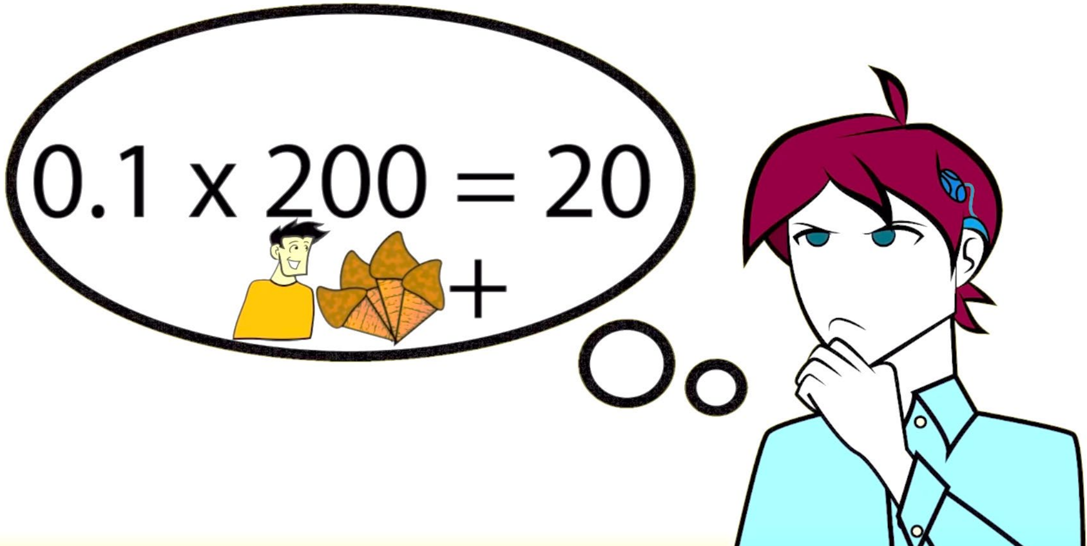
$$\begin{aligned} P(X>3) &= P(X=4)+P(X=5)+P(X=6) \\ &= 0.04 + 0.04 + 0.02 \\ &= 0.1 \text{ or } 10\% \end{aligned}$$

Out of

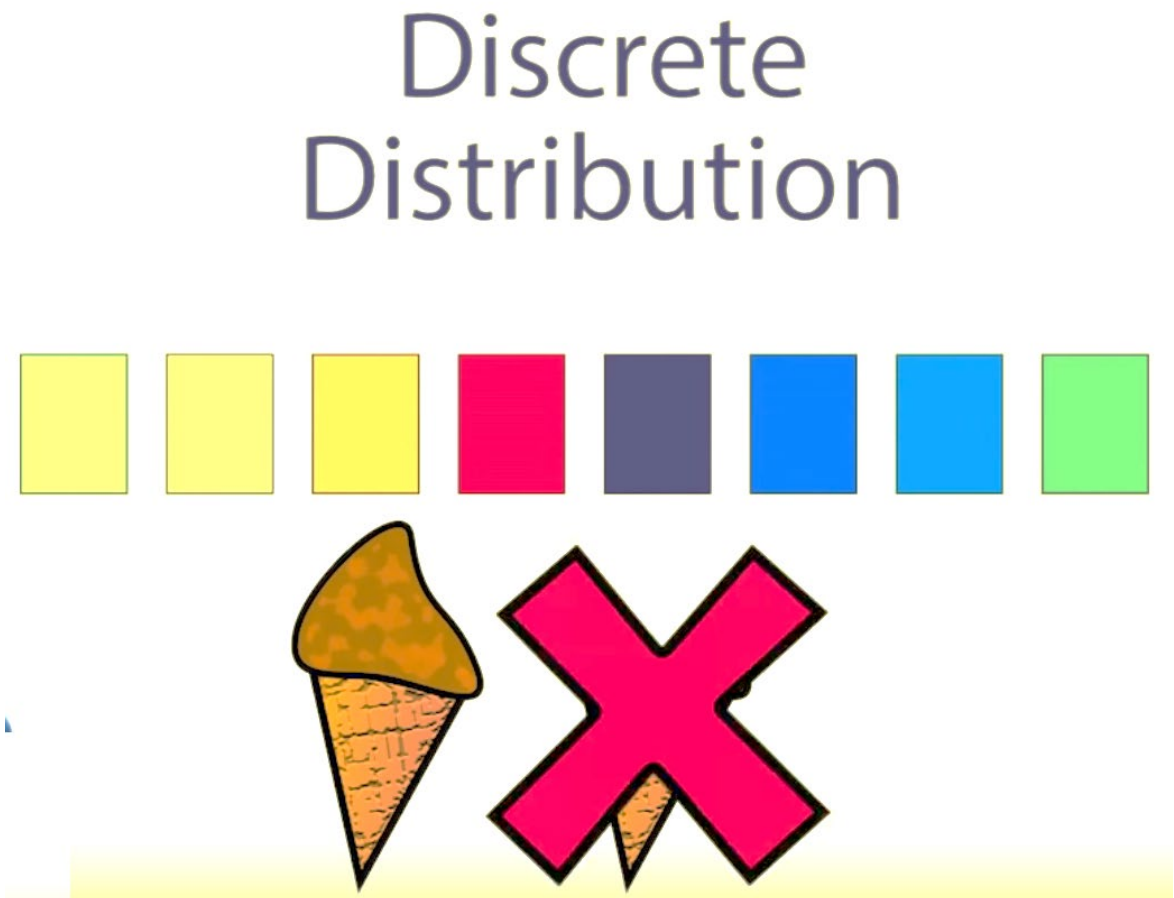
200

More than

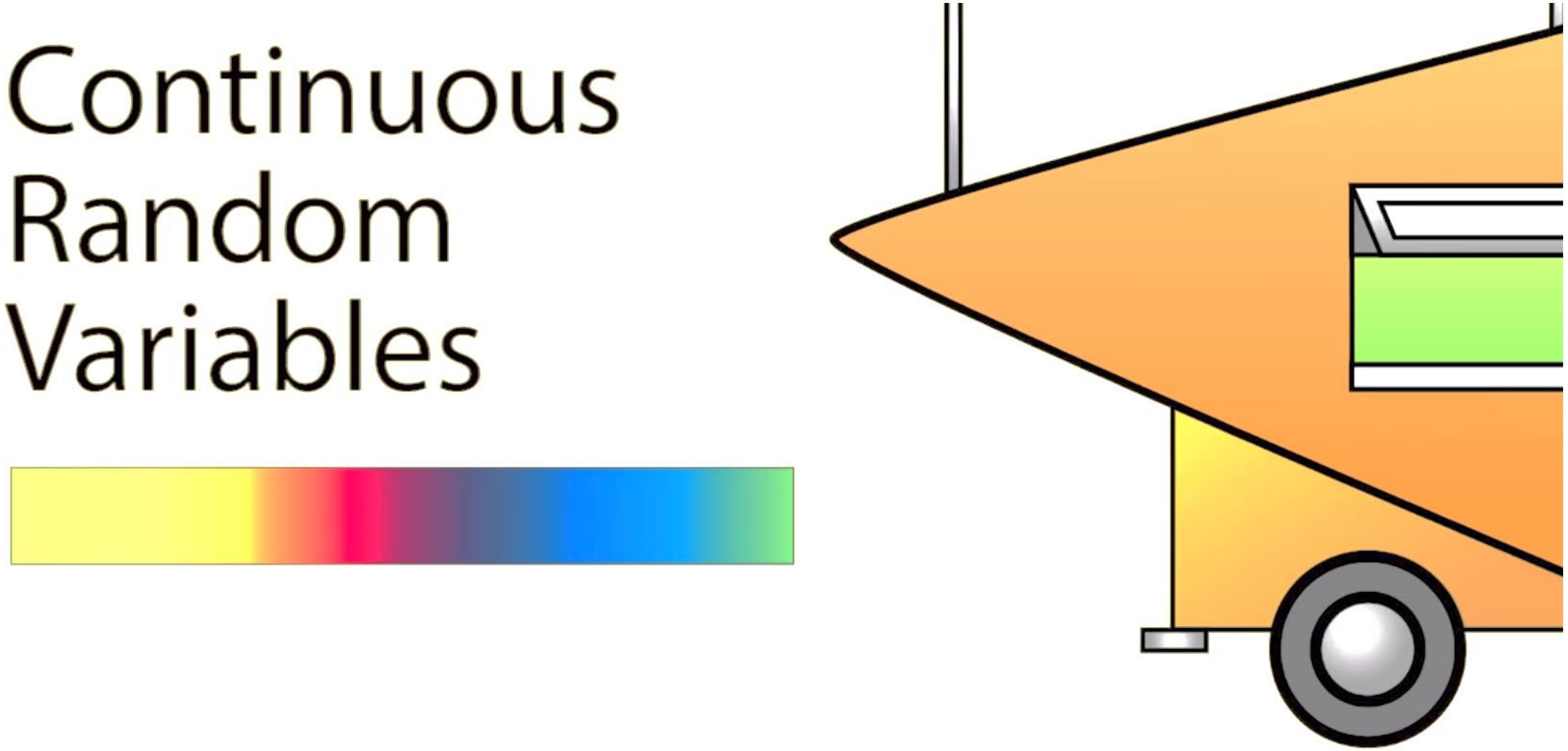




These are discrete distributions of R.V.

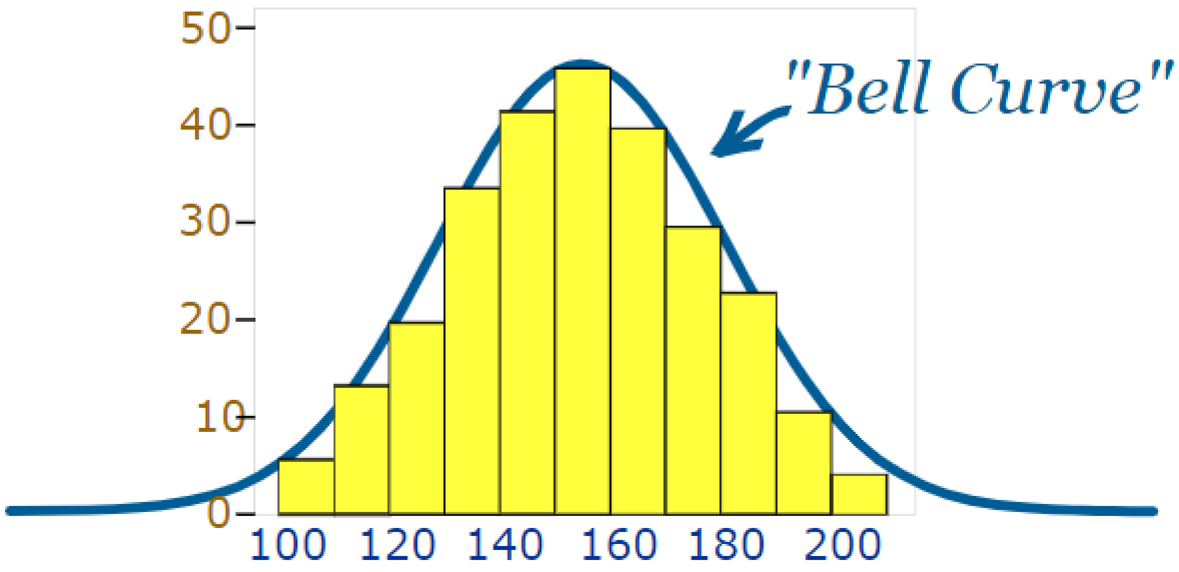


Discrete vs. Continuous Random Variables

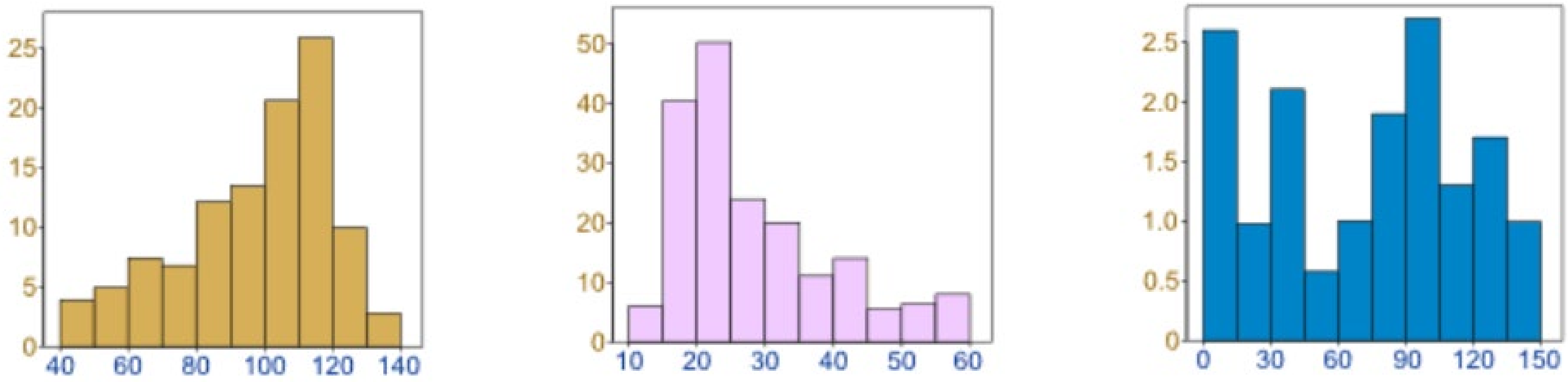


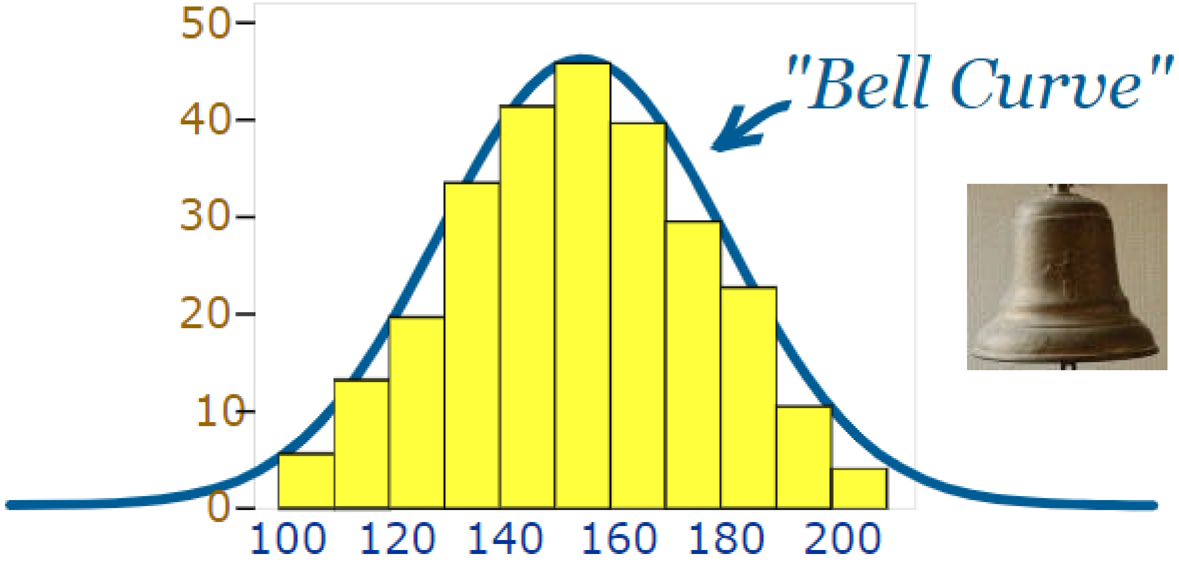
In practice, when an expected value would make sense to represent the entire distribution?

Question 1



Normal or “Bell” Distribution – A V.I.Distribution





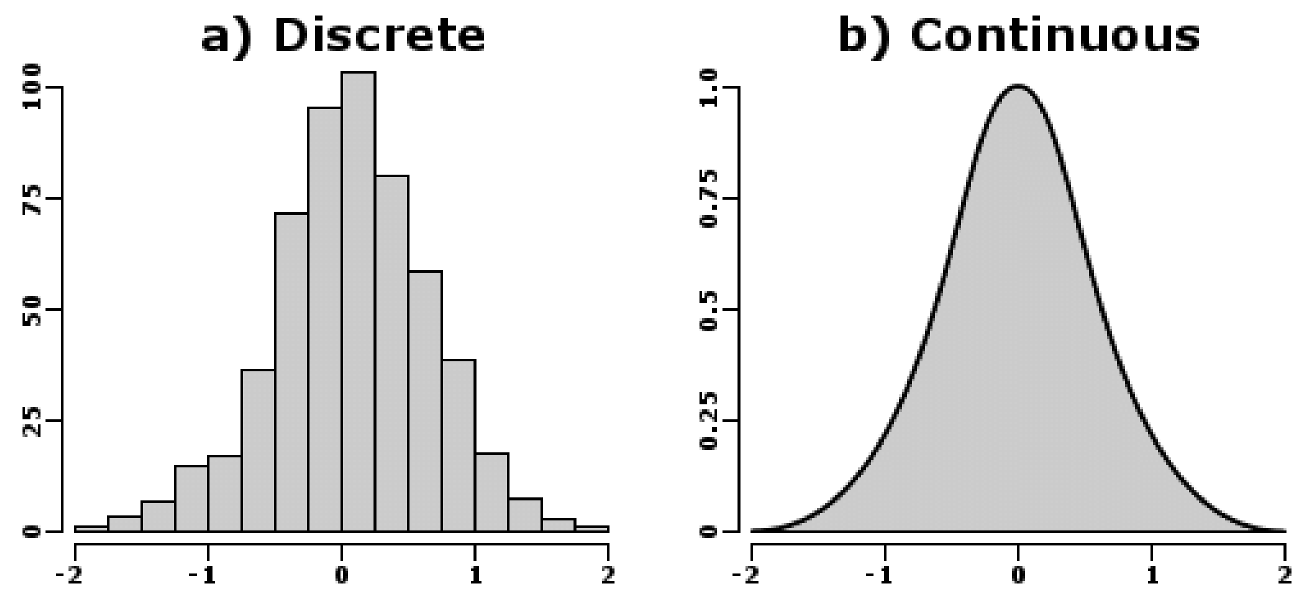
Many things closely follow a Normal Distribution:

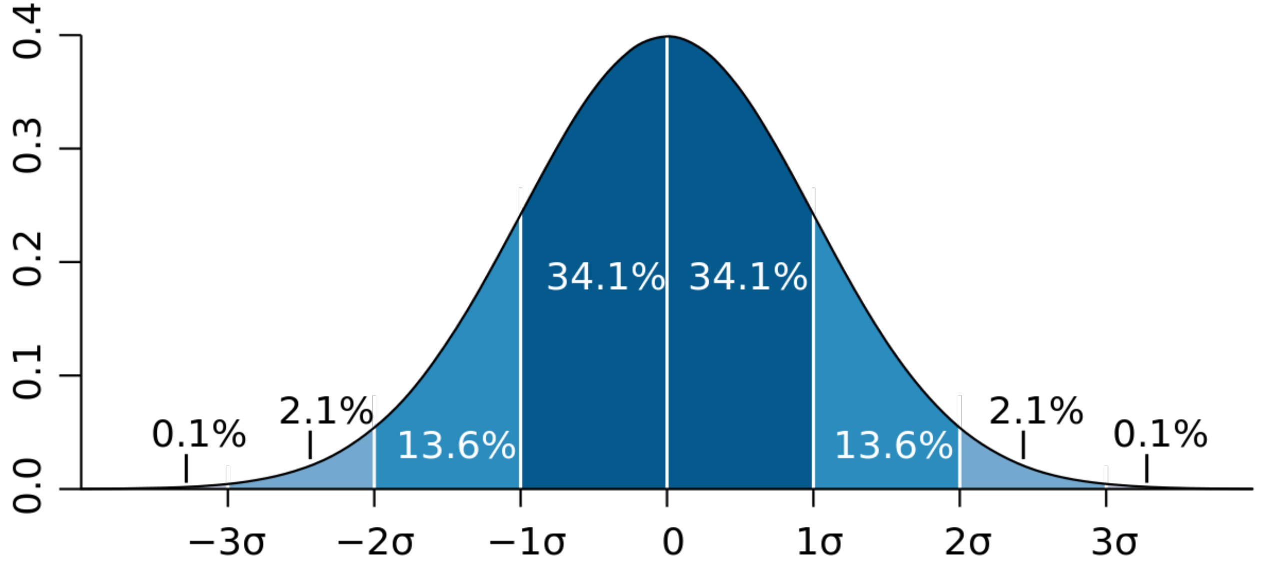
- size of things produced by machines
- errors in measurements
- Most physiological variables

– e.g., blood pressure, height, brain signals

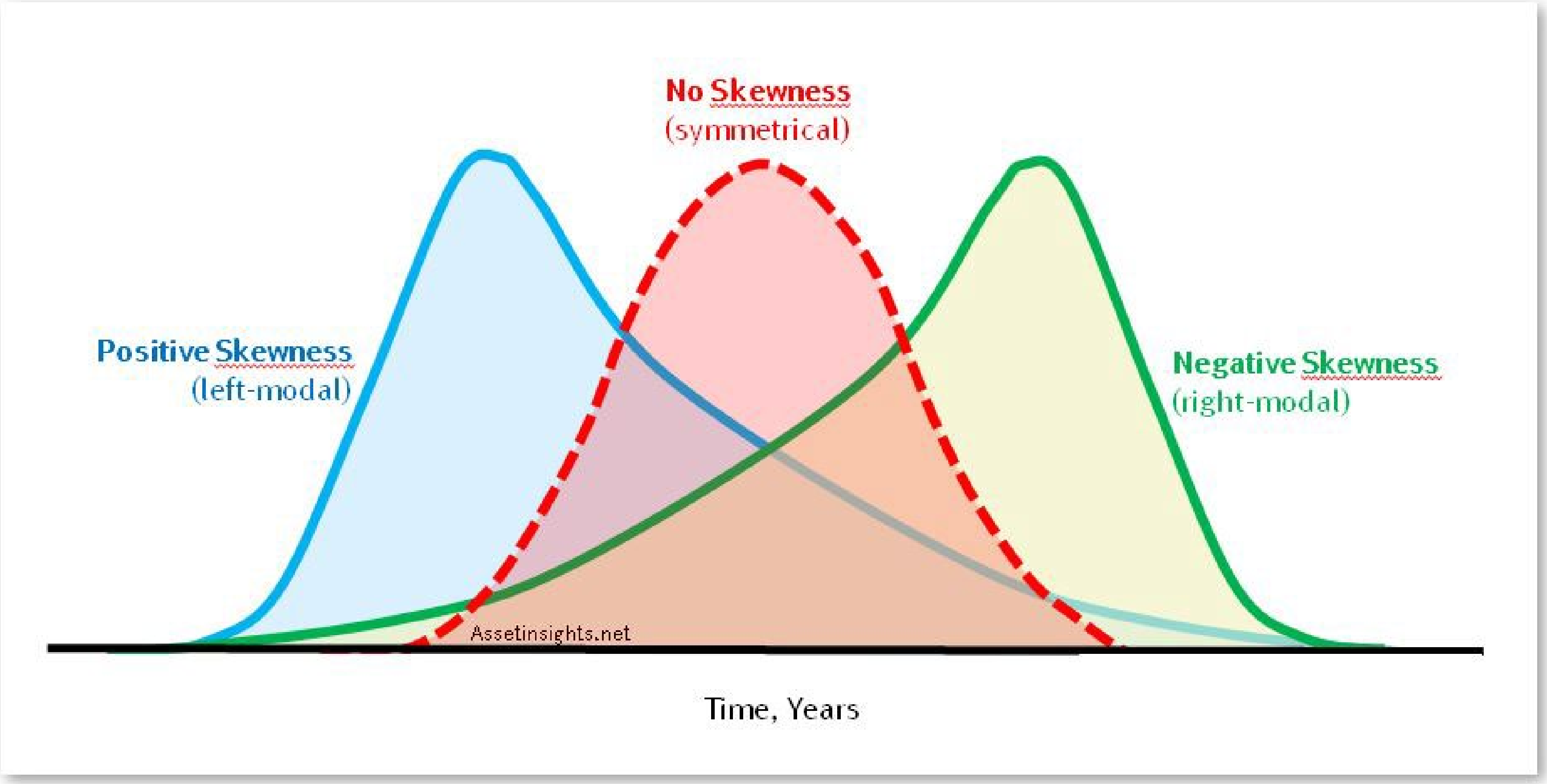
- marks on a test

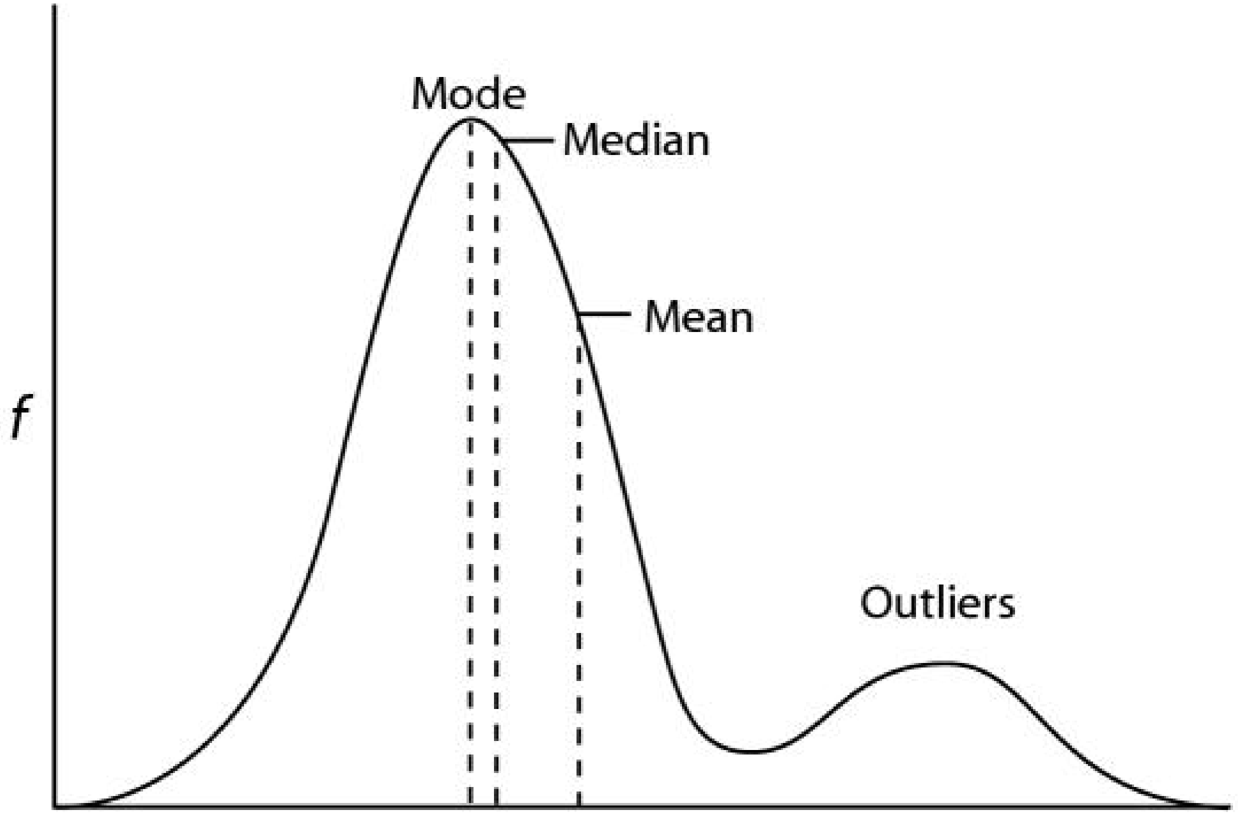
(that is one way to detect anomalies in your Quizzes)


a) Discrete
b) Continuous
100
75
50
25
0
1.0
0.75
0.5
0.25
0
-2
-1
0
1
2


0.4
0.3
0.2
0.1
0.0
34.1%
34.1%
13.6%
13.6%
2.1%
2.1%
0.1%
0.1%
−3σ
−2σ
−1σ
0
1σ
2σ
3σ

(Some) Types of Distributions





Mode
Median
Mean
f
Outliers

Probability Distribution – Mean or Expected Value

$$\mu = \Sigma xp$$

Example: Tossing a single **unfair** die

For fun, imagine a **weighted** die (cheating!) so we have these probabilities:

1	2	3	4	5	6
0.1	**0.1**	**0.1**	**0.1**	**0.1**	**0.5**

Example continued:

x	1	2	3	4	5	6
p	0.1	0.1	0.1	0.1	0.1	0.5
xp	0.1	0.2	0.3	0.4	0.5	3

$$\mu = \Sigma xp = 0.1+0.2+0.3+0.4+0.5+3 = 4.5$$

The expected value is 4.5

To calculate the Expected Value:

- multiply each value by its probability
- sum them up