Discrete & Continous Distributions

Mixtures

& Expectations

The Kolmogorov Axioms of probability theory

1) $\Pr(\Omega) = 1$

2) If V is a countable collection of disjoint events:

$$V = \{A_1, A_2, \ldots\}, \quad \forall i \neq j, \; A_i \cap A_j = \varnothing$$

Then Probability of the union is equal to the sum of the probabilities:

$$\Pr\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Pr(A_i)$$

The uniform distribution over [0,1]

$$0 \le x_1, x_2, \ldots \le 1$$

$$P(\{x_1\}) = 0$$

$$P(\{x_1, x_2\}) = 0$$

$$P(\{x_1, x_2, x_3, \ldots\}) = 0$$

But, if $0 \le a < b \le 1$

$$P([a,b]) = b - a$$

Prob of other sets:

Construct from Countable Unions & Intersections of intervals

U(A,B) = The Uniform distribution over the segment [A,B]

$U(A,B)$ is defined by assigning probability

to every segment $[a,b]$ where $A \le a \le b \le B$ (A<B)

$$\Pr\left([a,b]\right) = \Pr\left((a,b)\right) = \frac{b-a}{B-A}$$

Lets calculate the probability of some sets with respect to the uniform distribution

Fix the probability distribution $U(-1,1)$

$P([-1/3,1/3])=(1/3--1/3)/(1--1)=(2/3)/2=1/3$

$P([-1,0])=$

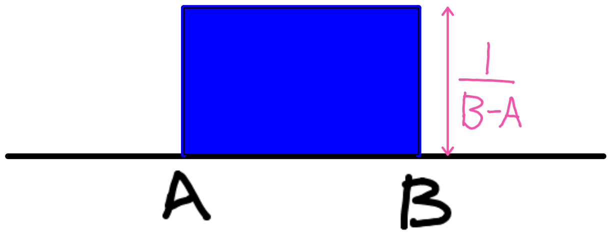
$P([-2,0])=$

$P([-3,2])=$

$P([0,2])=$

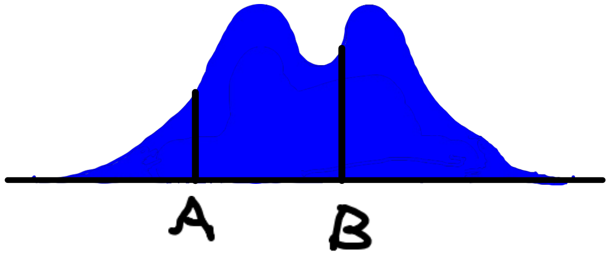
$P([-2,-1/2]U[1/2,2])=$

uniform density:



general density

$$P([A,B]) = \int_A^B f(x)\,dx$$



$\forall x \quad f(x) \geq 0 \qquad$ ($f(x)$ can be larger than 1)

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

PDF and CDF

The Probability Density function is shortened to PDF

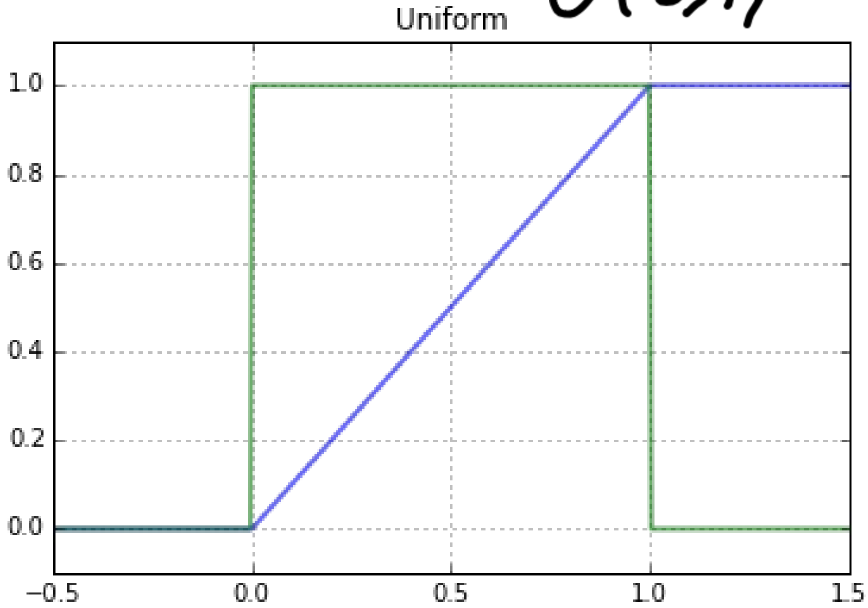
Another popular representation of a distribution on the real is the CDF: Cumulative Distribution function

The CDF F is defined as $F(a) \doteq \Pr(x \leq a)$

For density distributions, one can translate between PDF and CDF:

$$F(a) = \int_{-\infty}^{a} f(x)dx; \quad f(a) = \left.\frac{dF(x)}{dx}\right|_{x=a}$$

$U(0,1)$



Uniform Dist. $\boldsymbol{U(a, b)}$

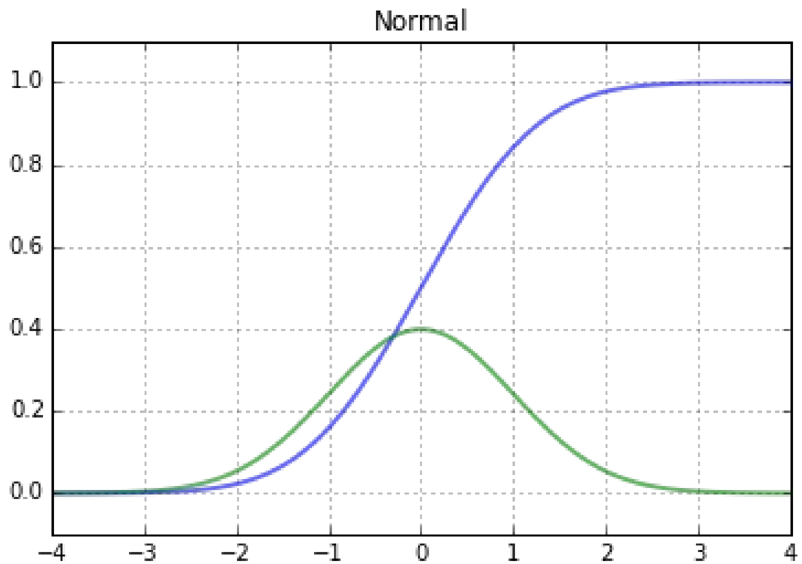
Endpoints: $a = 0,\ b = 1$

PDF:

$$f(x) = \begin{cases} 0 & x < a \\ \frac{1}{b-a} & a \le x < b \\ 0 & b \le x \end{cases}$$

CDF:

$$F(x) = \int_{-\infty}^{x} f(s)ds = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x < b \\ 1 & b \le x \end{cases}$$



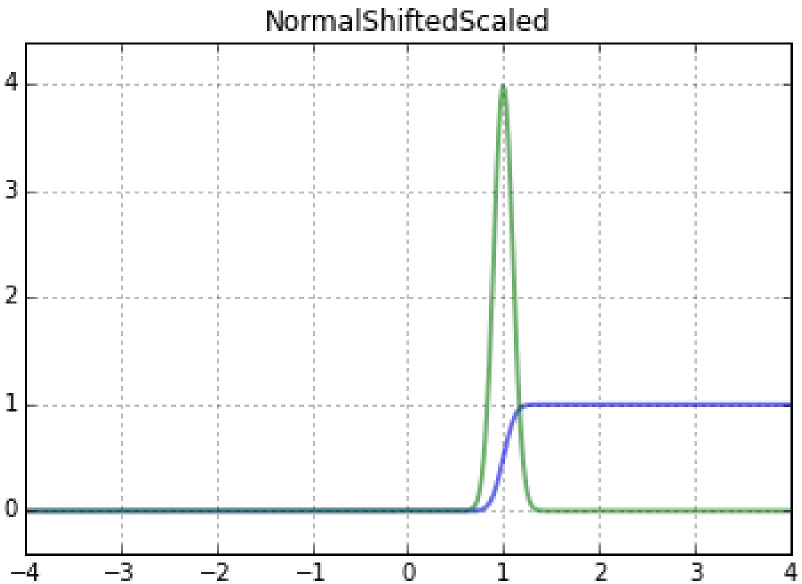
Standard Normal N(0,1)

PDF:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

CDF:

$$F(x) = \int_{-\infty}^{x} f(s) ds = 1 - Q(x)$$



Shifted and Scaled Normal $\boldsymbol{N(\mu, \sigma)}$

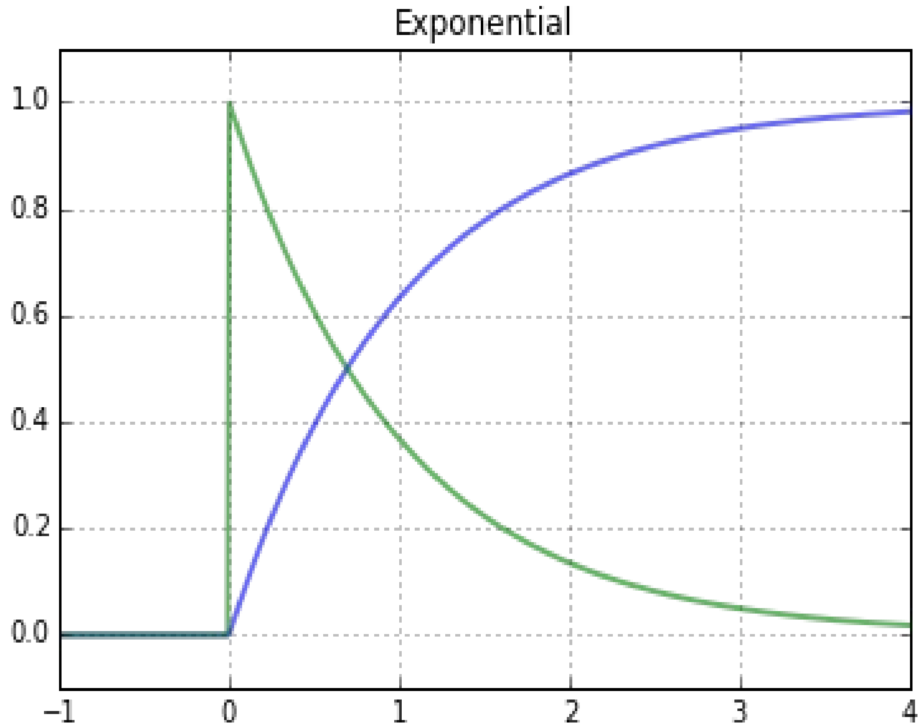
Shift: $\mu = 1$ *scale:* $\sigma = 0.1$

PDF:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

CDF:

$$F(x) = \int_{-\infty}^{x} f(s)ds = 1 - Q\left(\frac{x-\mu}{\sigma}\right)$$



Exponential Distribution $\color{red}{\boldsymbol{Exp(a,\lambda)}}$

Shift: a *Scale:* $\lambda > 0$

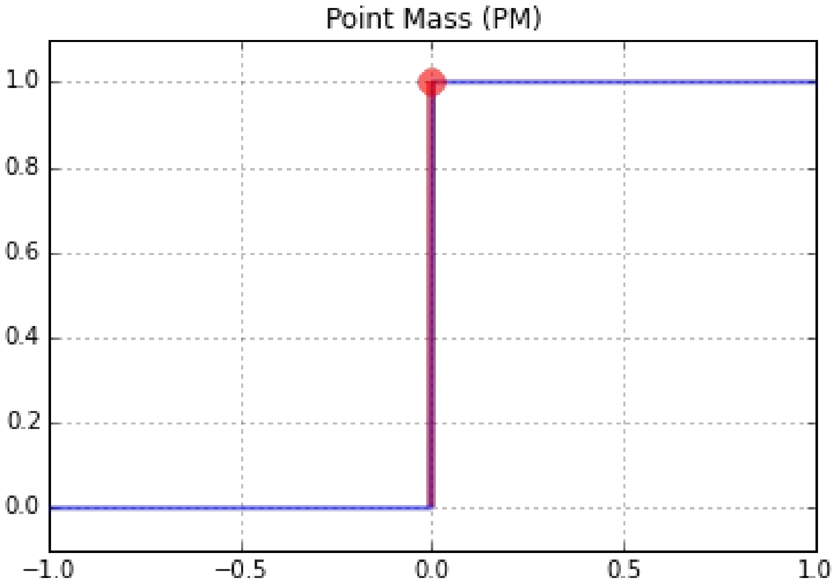
PDF:

$$f(x) = \begin{cases} 0, & x < a \\ \lambda e^{-\lambda(x-a)}, & x \geq a \end{cases}$$

CDF:

$$F(x) = \begin{cases} 0, & x < a \\ 1 - e^{-\lambda(x-a)}, & x \geq a \end{cases}$$

PM(0)



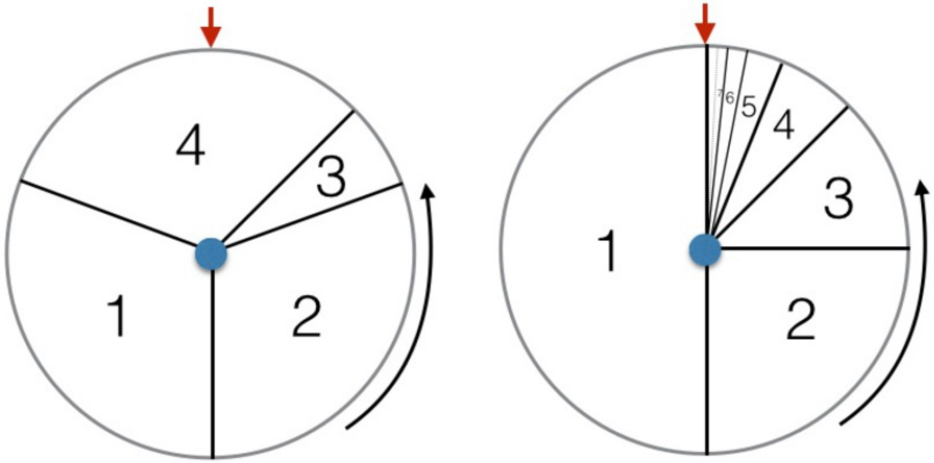
Point-mass distribution PM(a)
Shift: a

PMF:

$$P(a) = 1$$

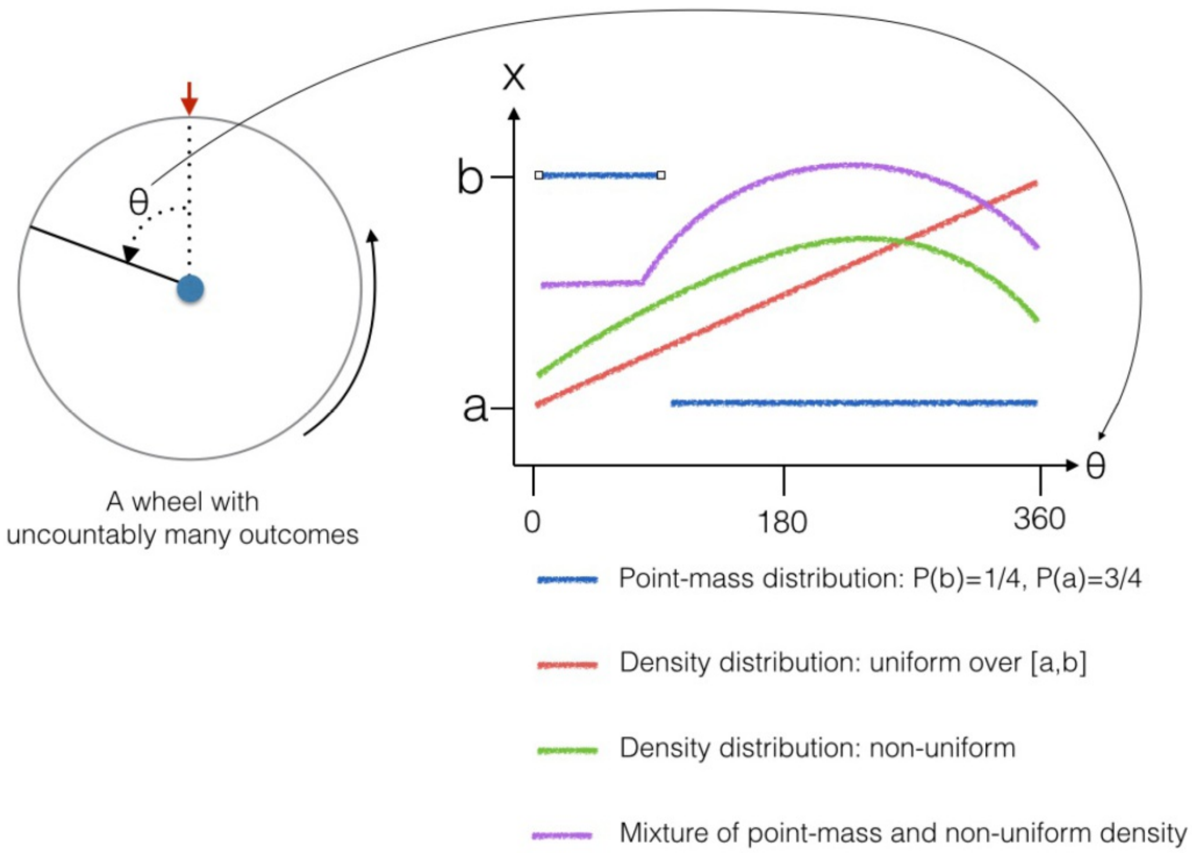
CDF:

$$F(x) = \int_{-\infty}^{x} f(s)ds = \begin{cases} 0 & x < a \\ 1 & x \geq a \end{cases}$$



(a) A wheel with four outcomes

(b) A wheel with Infinitely many outcomes


θ
A wheel with
uncountably many outcomes
X
b
a
θ
0
180
360
Point-mass distribution: P(b)=1/4, P(a)=3/4
Density distribution: uniform over [a,b]
Density distribution: non-uniform
Mixture of point-mass and non-uniform density

density distributions vs. Point-Mass distribution

Point mass distributions assign non-zero probability to individual points.
PM(a) ---- P(X=a)=1

Density distributions assign non-zero probability to segments.

The probability of any single point under a density distribution is zero.
=> as a result P([a,b])=P((a,b))=P([a,b))=P((a,b])

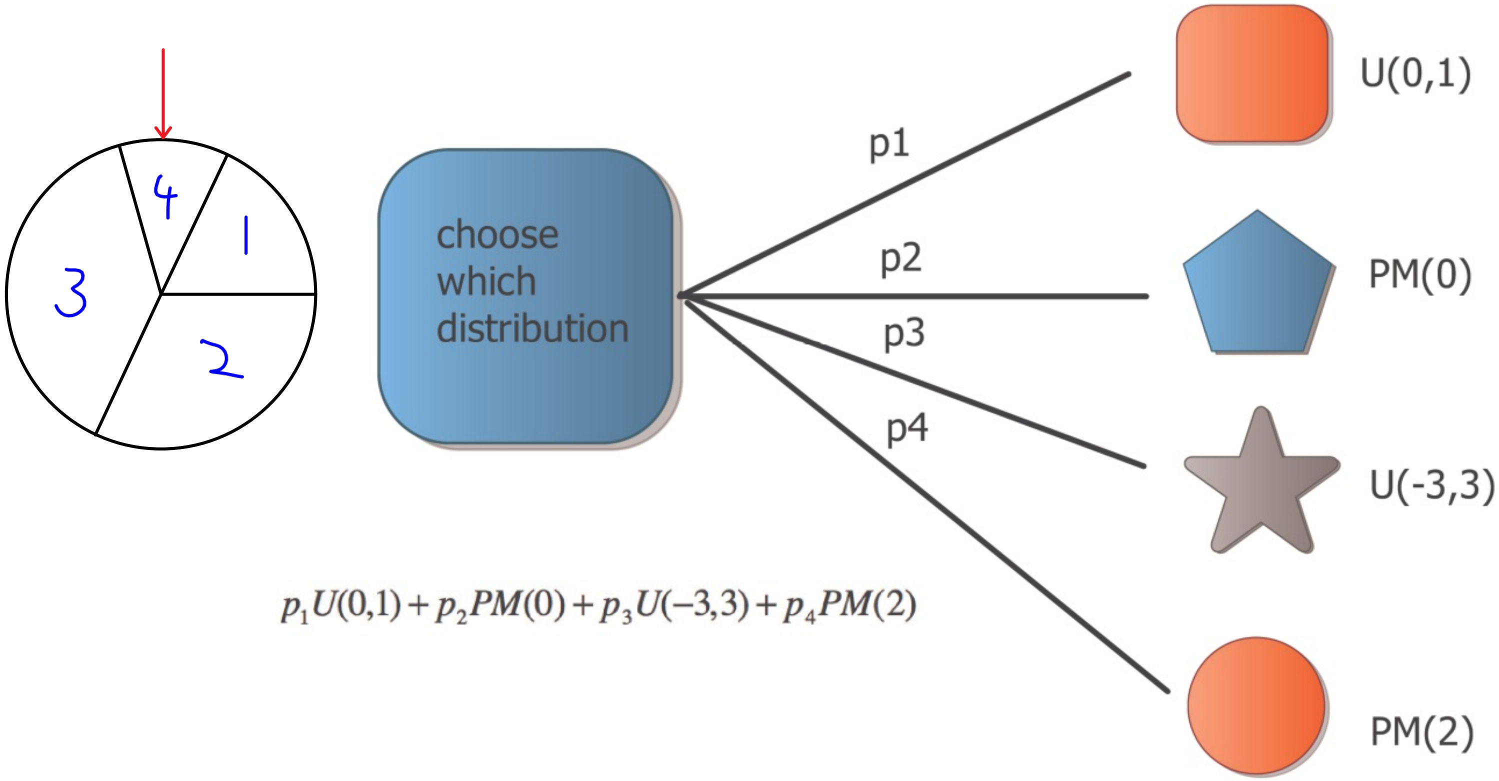
=> the probability of any countable set is zero.

=> for example the probability of all rational numbers in [0,1], under the uniform distribution over [0,1] is zero!!!

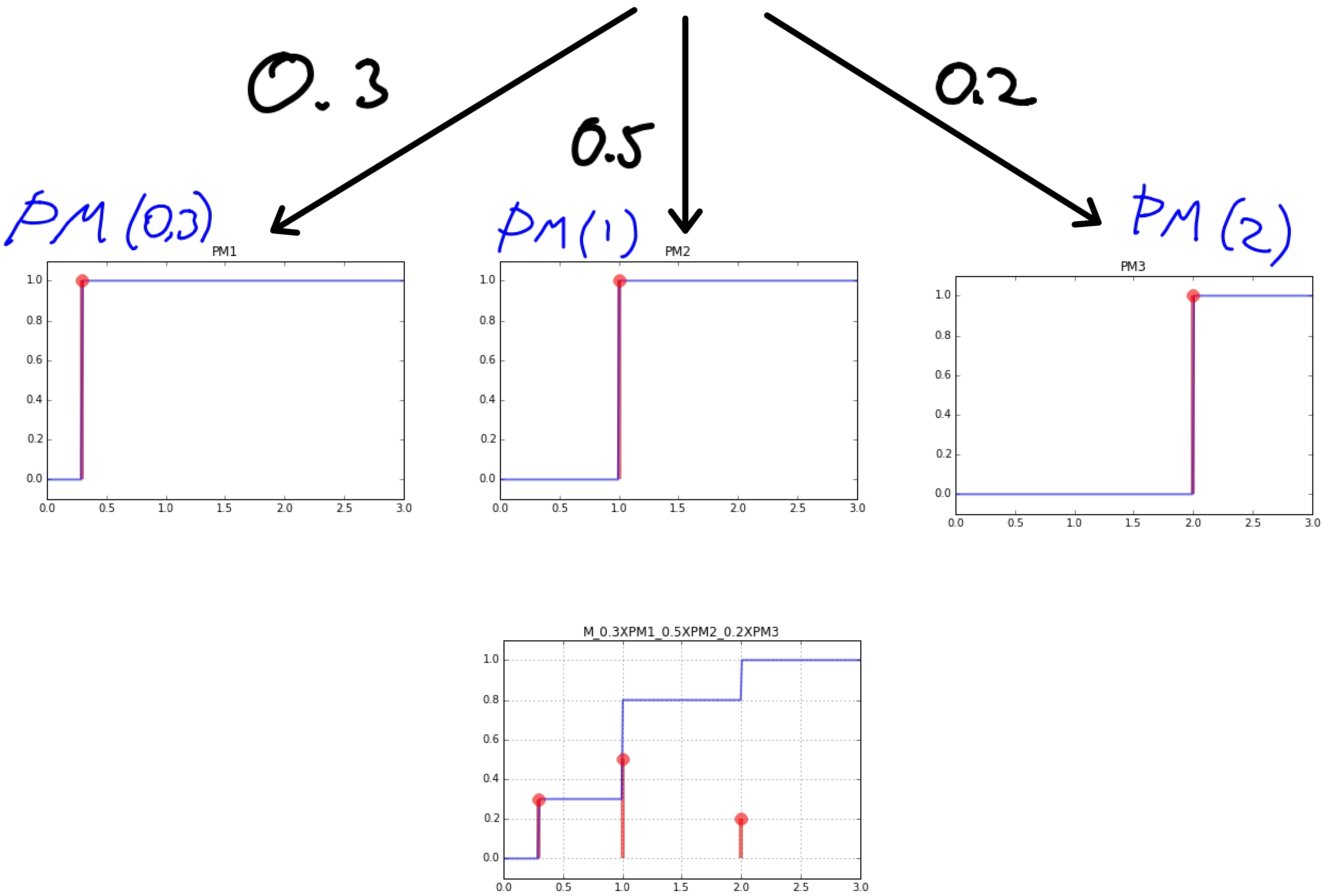
In other words, if you pick a random number from U(0,1)
the probability that it is a rational number is zero !!!

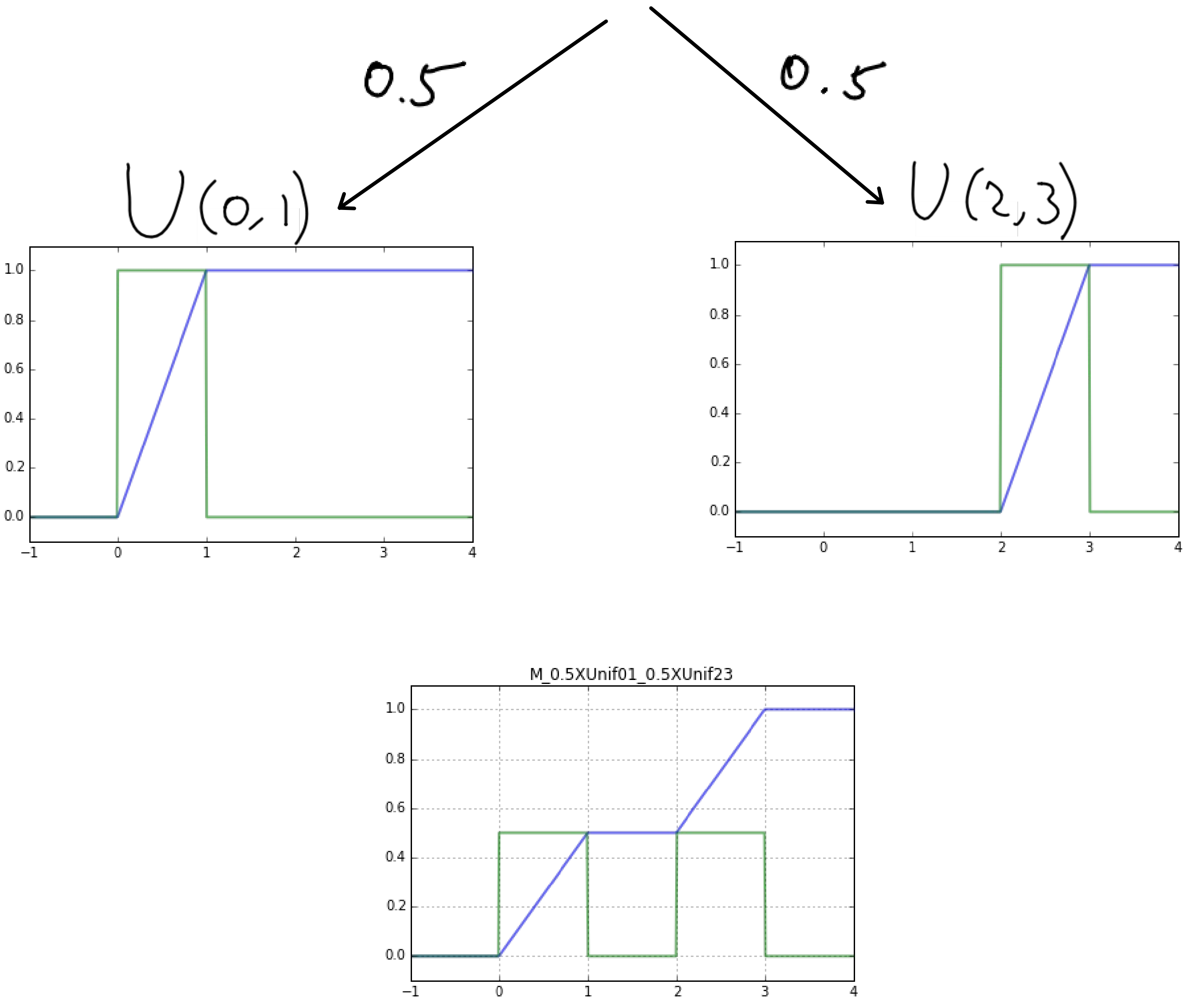
Mixture Distributions

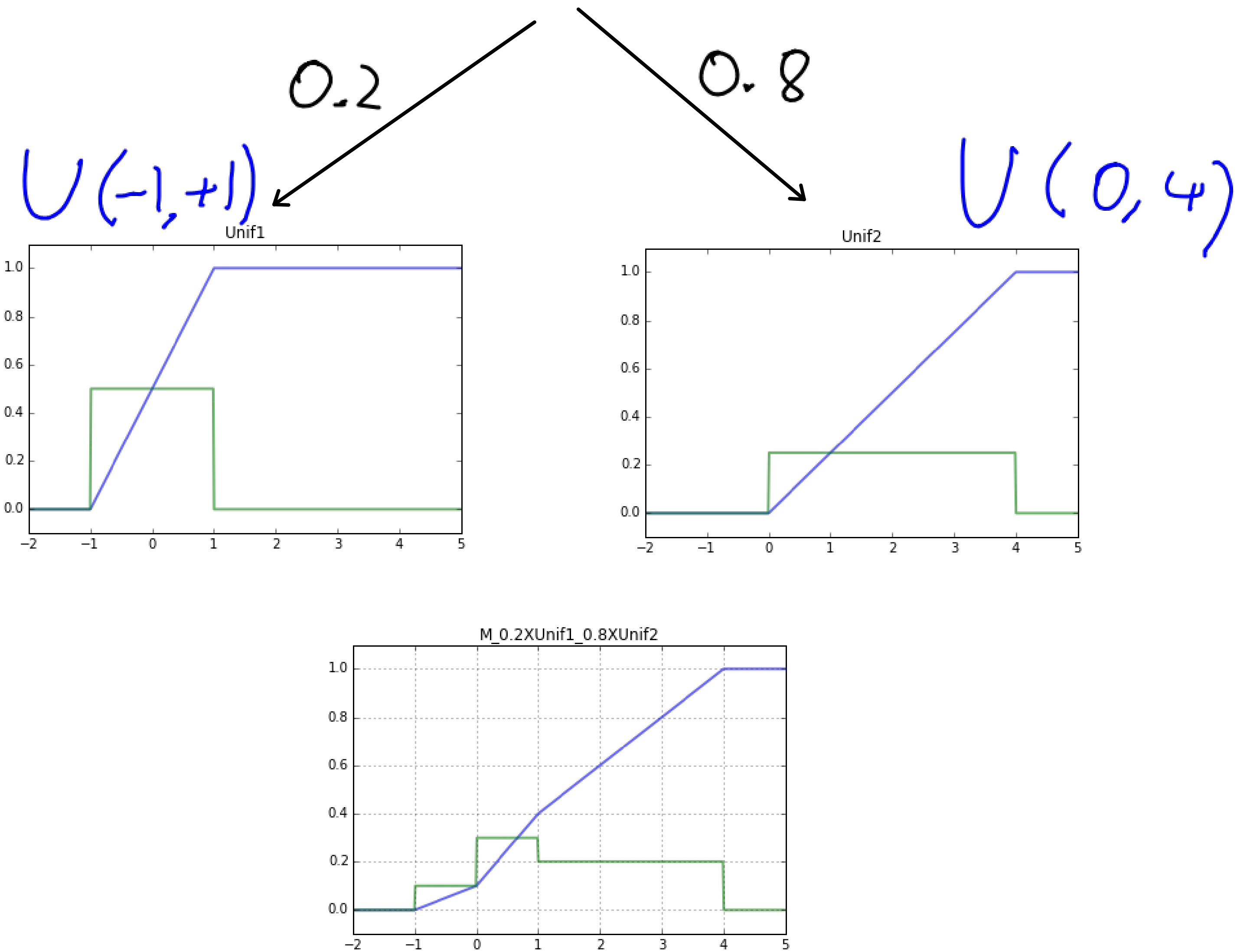
Mixtures distributions

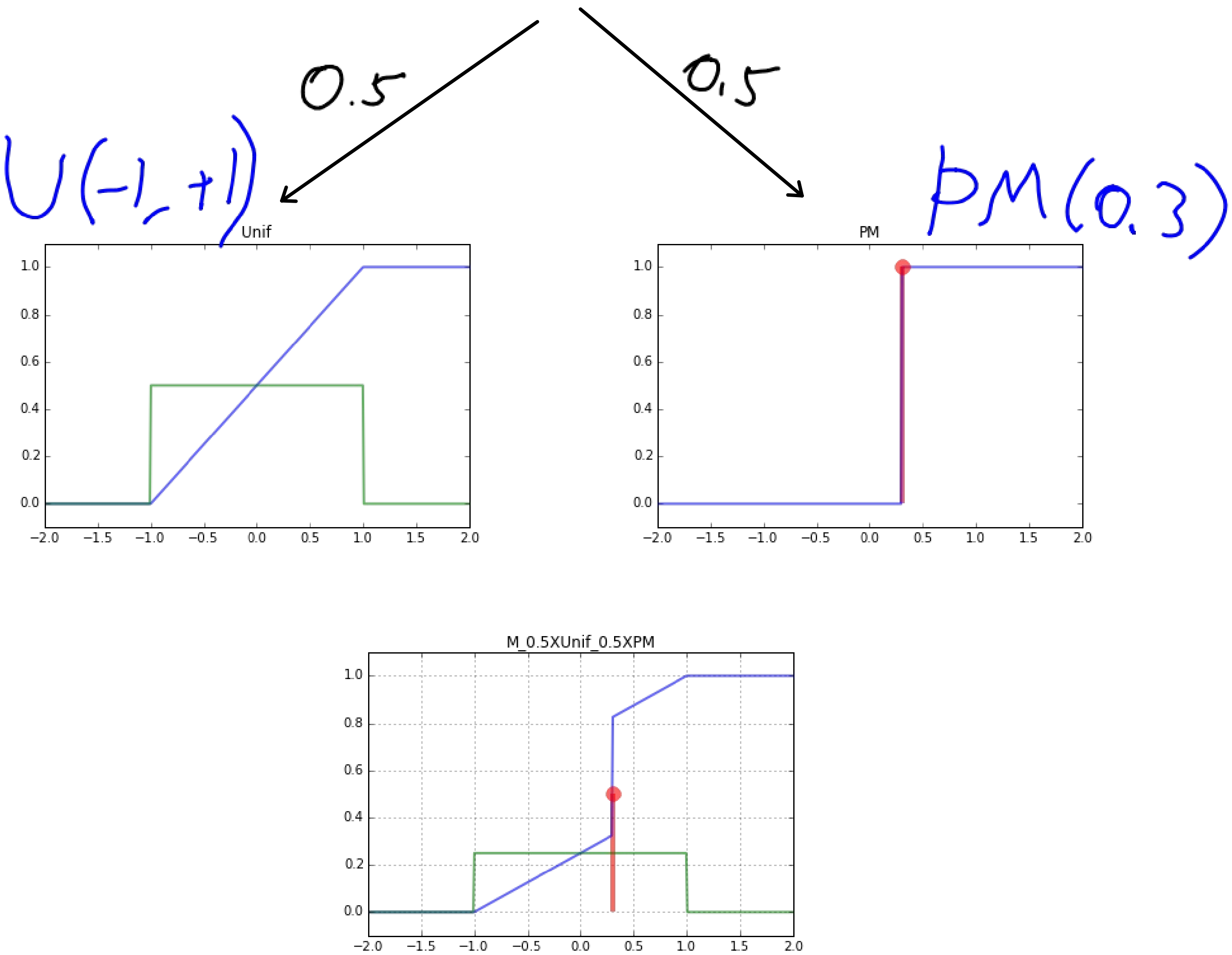


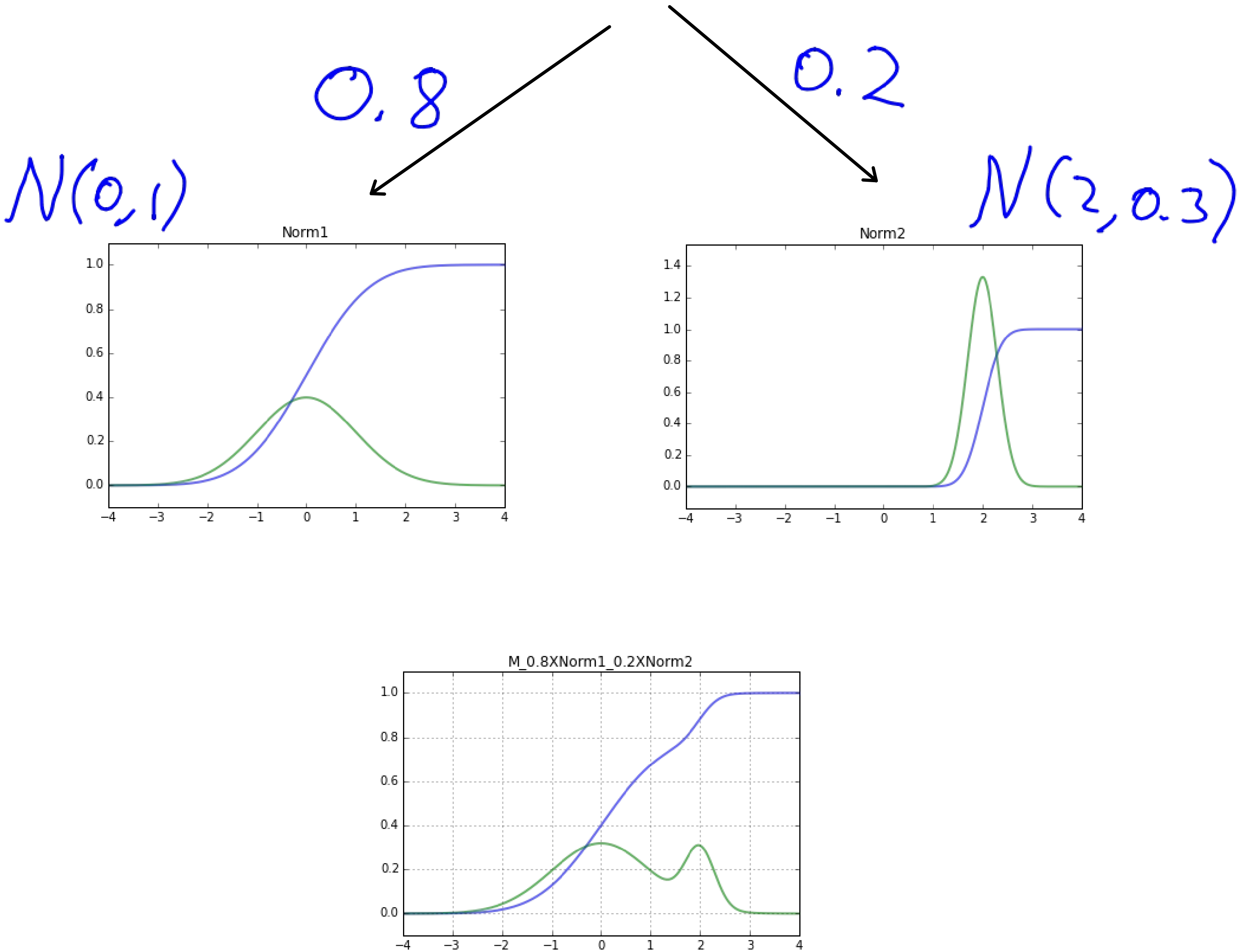
$$p_1 U(0,1) + p_2 PM(0) + p_3 U(-3,3) + p_4 PM(2)$$


0.3
0.5
0.2
PM(0.3)
PM(1)
PM(2)
PM1
PM2
PM3
M_0.3XPM1_0.5XPM2_0.2XPM3

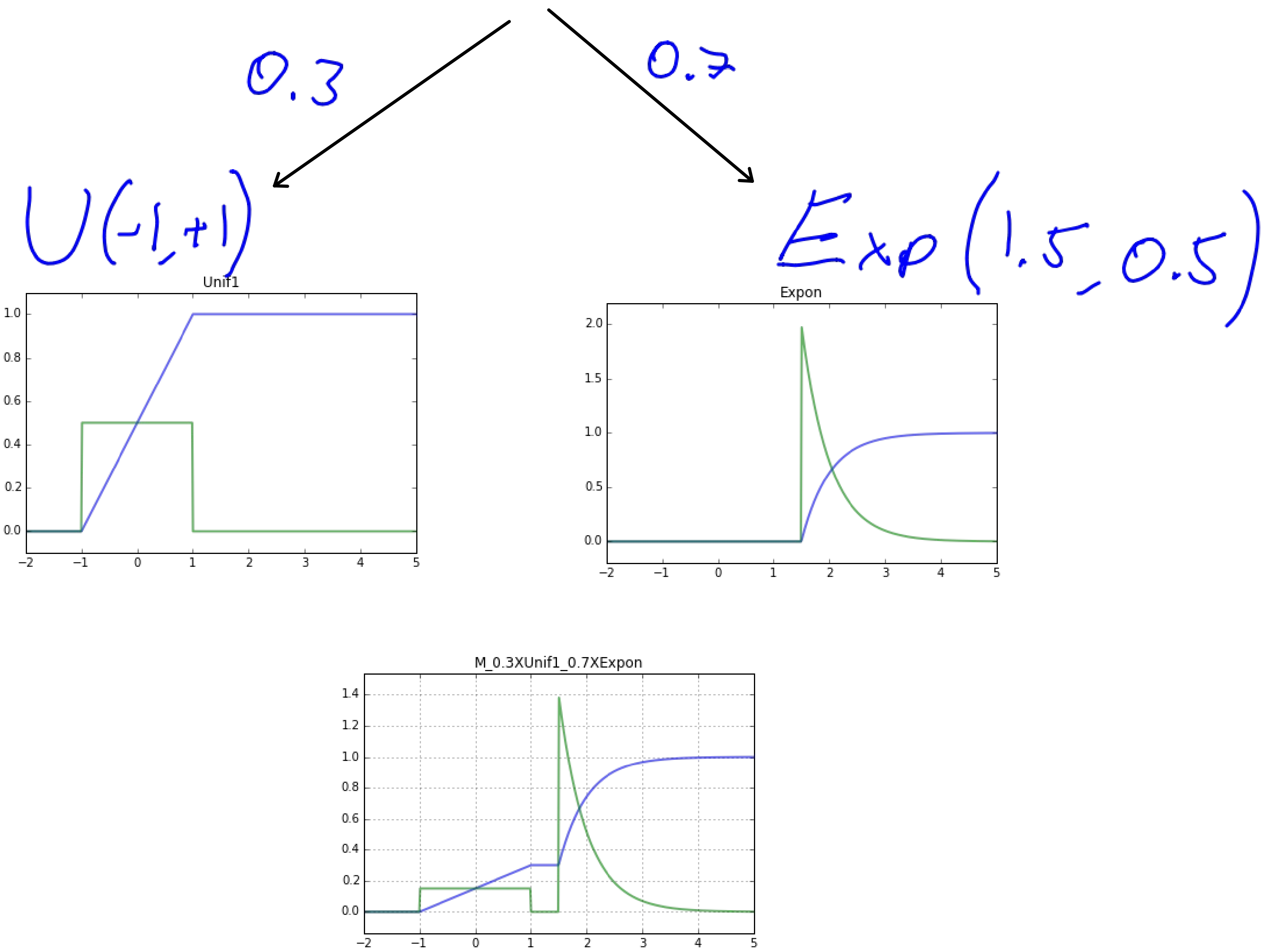

0.5
0.5
U(0,1)
U(2,3)
M_0.5XUnif01_0.5XUnif23


0.2
0.8
U(-1,+1)
U(0,4)
Unif1
Unif2
M_0.2XUnif1_0.8XUnif2


0.5
0.5
U(-1, +1)
PM(0.3)
Unif
PM
M_0.5XUnif_0.5XPM



0.8
0.2
N(0,1)
N(2,0.3)
Norm1
Norm2
M_0.8XNorm1_0.2XNorm2



0.3
0.7
U(-1,+1)
Exp(1.5, 0.5)
Unif1
Expon
M_0.3XUnif1_0.7XExpon

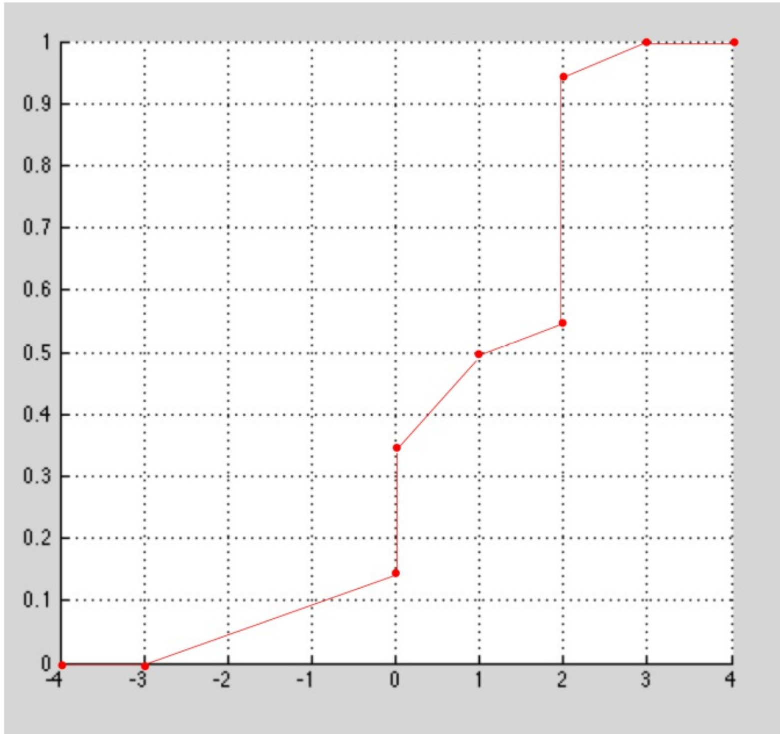
$$.1U(0,1) + .2PM(0) + .3U(-3,3) + .4PM(2)$$

$$F(-3) = 0;\ F(-.01) \approx .5 * .3 = .15$$

$$F(0) = .35;\ F(1) = .35 + .1 + \frac{.3}{6} = 0.5;$$

$$F(1.99) \approx 0.5 + 0.05 = 0.55;\quad F(2) = 0.95$$

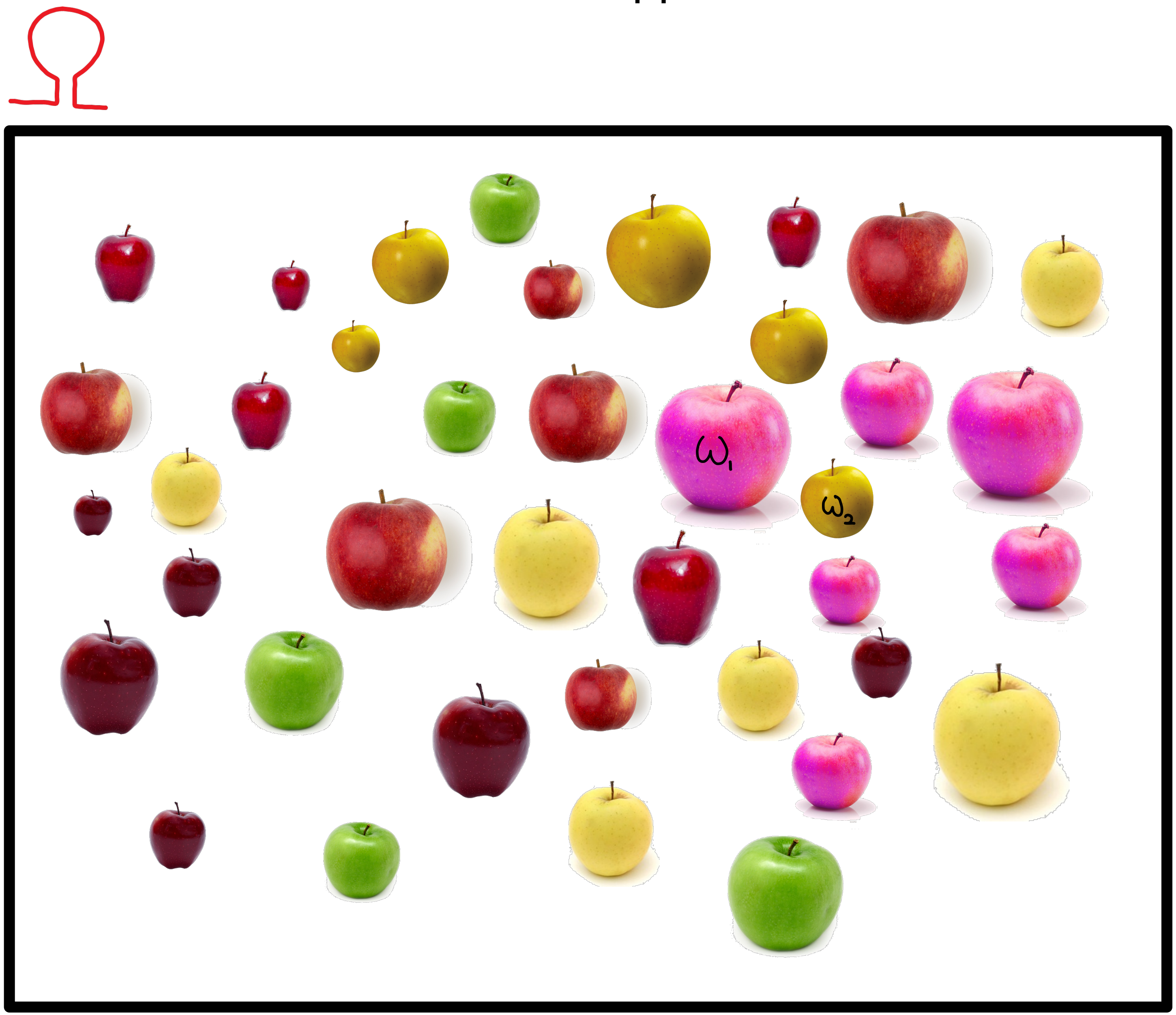
$$F(3) = 0.95 + \frac{.3}{6} = 1.0$$

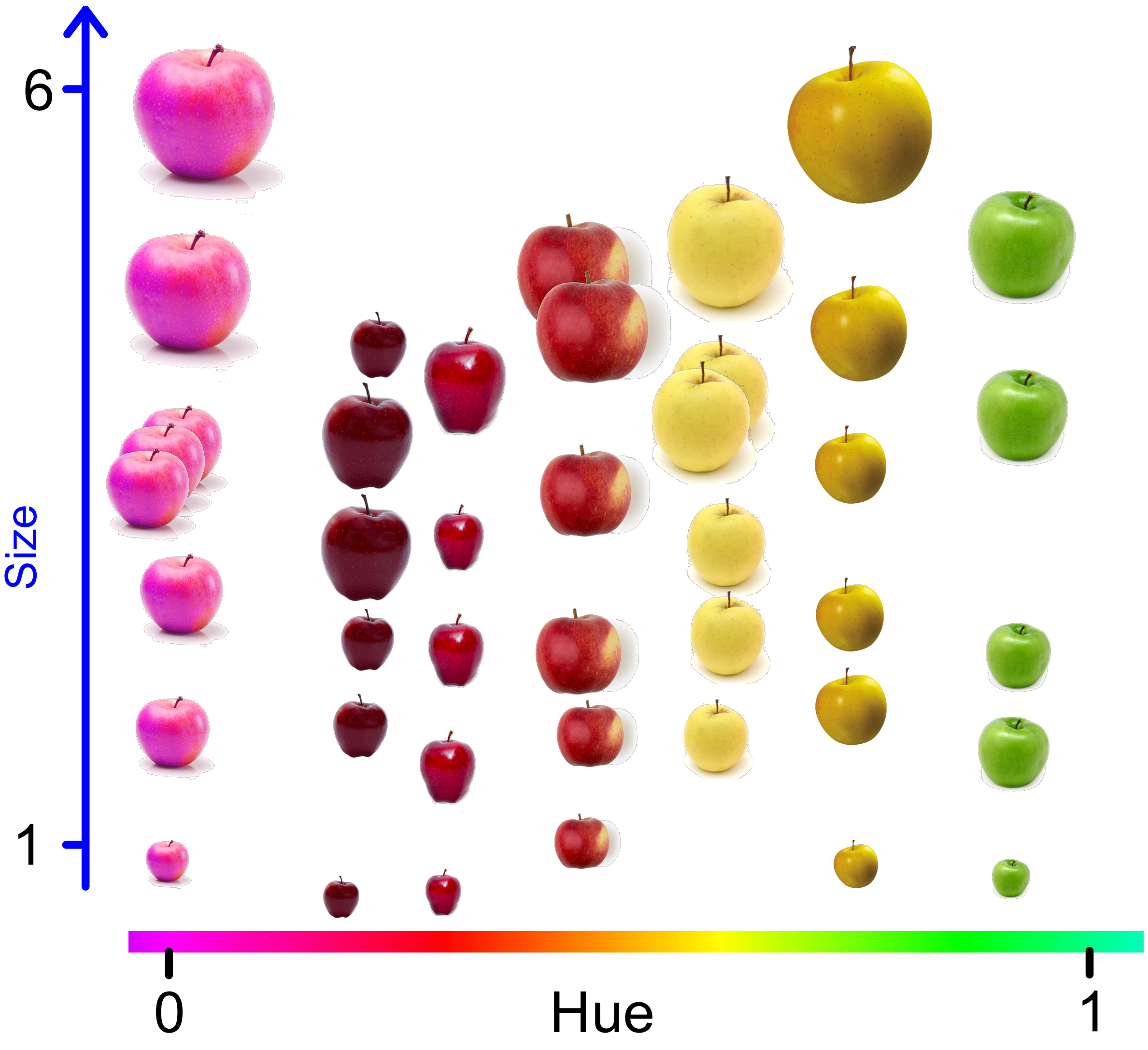


Random Variables

(RVs)

Sample space = apples
An outcome is an apple





Size	Hue

baseball acronyms

G Games Played
PA Plate Appearances
AB At Bat
R Runs Scored
H Hits
D ?
T ?
HR Home Runs
RBI Runs Batted In
BB Bases on Balls (walks)
SO Strikeouts
BA ?
OBP On base Percentage
SLG Slugging Percentage
OPS OBP+SLG

Player	G	PA	AB	R	H	D	T	HR	RBI	BB	SO	BA	OBP	SLG	OPS
Mike Napoli	17	64	49	13	16	4	0	6	14	15	16	.327	.484	.776	1.260
Josh Donaldson	20	88	72	17	28	8	0	5	15	14	11	.389	.500	.708	1.208
Hunter Pence	21	92	78	19	25	4	0	9	26	13	14	.321	.413	.718	1.131
Matt Carpenter	21	102	88	23	35	11	2	1	11	12	19	.398	.480	.602	1.083
Ryan Zimmerman	21	97	90	22	29	2	0	**11**	16	7	18	.322	.371	.711	1.082
Freddie Freeman	20	85	73	15	26	3	0	6	17	10	15	.356	.435	.644	1.079
Michael Cuddyer	16	67	62	8	26	4	0	3	14	4	10	.419	.448	.629	1.077
Adam Lind	18	60	55	11	16	2	0	7	17	5	11	.291	.350	.709	1.059
Andrew McCutchen	20	83	66	13	22	5	1	3	8	13	11	.333	.470	.576	1.046
Prince Fielder	20	86	79	11	30	7	0	4	14	6	13	.380	.419	.620	1.039
Shin-Soo Choo	18	85	60	15	18	3	0	4	10	21	12	.300	.488	.550	1.038
Paul Goldschmidt	21	92	80	12	27	6	2	4	19	11	19	.338	.424	.613	1.036
Moises Sierra	19	67	63	8	22	12	1	1	9	4	14	.349	.388	.619	1.007
Josmil Pinto	16	62	58	9	21	5	0	3	9	4	12	.362	.403	.603	1.007
Mike Trout	21	95	71	16	21	5	1	3	10	23	21	.296	.474	.521	.995
Yoenis Cespedes	19	80	77	12	26	2	1	6	19	2	19	.338	.363	.623	.986
Matt Holliday	20	92	76	14	28	6	0	2	20	14	13	.368	.457	.526	.983
David Ortiz	20	91	76	18	21	9	0	5	16	13	16	.276	.385	.592	.977
Chase Headley	17	67	56	8	15	2	0	5	9	10	12	.268	.388	.571	.959
Matt Adams	19	72	69	13	21	1	0	7	14	3	21	.304	.333	.623	.957
Joey Votto	20	94	73	12	22	3	0	4	9	20	17	.301	.447	.507	.954
Eric Hosmer	20	87	78	12	27	6	1	2	12	9	18	.346	.414	.526	.939
Wil Myers	20	84	78	10	24	9	0	4	10	6	18	.308	.357	.577	.934
Giancarlo Stanton	20	85	72	12	19	3	0	6	16	11	27	.264	.376	.556	.932
Desmond Jennings	21	83	68	7	19	6	1	3	13	13	16	.279	.398	.529	.927

Outcome space: ***all possible performances of baseball hitters for a month***
Outcome: ***The performance of a particular player***
Random variables: ***measures of performance: G, PA, AB ...***

Events: ***More than 8 home runs,***
OPS higher than 1.0, 1.1, 1.2, ...

Event & RVs

from RV to event

$$A = \{\omega \in \Omega \mid X(\omega) > 5\}$$

from Event to RV

$$X = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

RVs $X(\omega), Y(\omega)$ are independent

if $\boxed{\begin{array}{l} \forall A, B,\ A \text{ defined using } X \\ \qquad\quad B \text{ defined using } Y \\ \quad A, B \text{ are independent} \end{array}}$

Joint distribution of two independent random variables

	X=1	X=2	X=10	P(Y=y)
Y=-1	1/12	1/12	2/12	4/12= 1/3
Y=+1	2/12	2/12	4/12	8/12= 2/3
P(X=x)	3/12= 1/4	3/12= 1/4	6/12= 1/2	

Marginals

Joint distribution of two dependent random variables

	X=1	X=2	X=10	P(Y=y)
Y=-1	1/12	2/12	1/12	4/12= 1/3
Y=+1	2/12	1/12	5/12	8/12= 2/3
P(X=x)	3/12= 1/4	3/12= 1/4	6/12= 1/2	

Marginals

Expected Value

Expected Value

pages 104-106
In OpenIntro
statistics

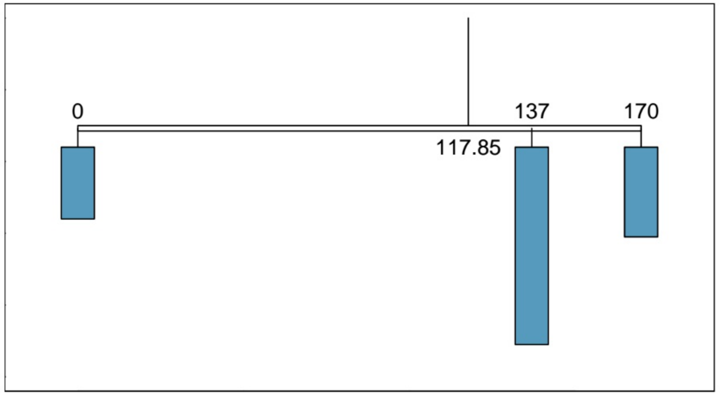


Figure 2.22: A weight system representing the probability distribution for X. The string holds the distribution at the mean to keep the system balanced.

i	1	2	3	Total
x_i	\$0	\$137	\$170	–
$P(X = x_i)$	0.20	0.55	0.25	1.00

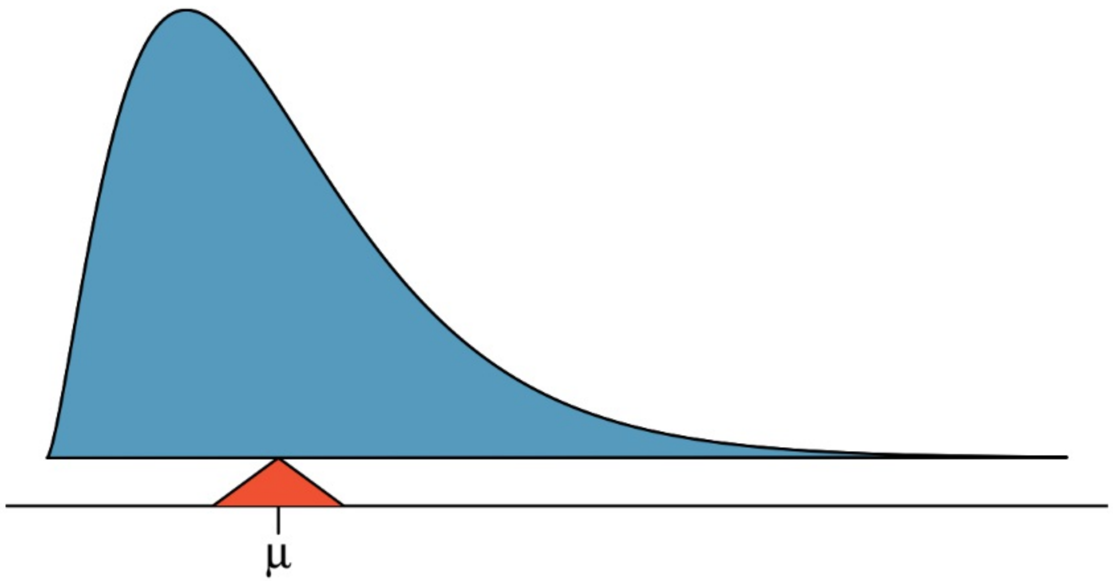
$$
\begin{aligned}
E(X) &= 0 \times P(X = 0) + 137 \times P(X = 137) + 170 \times P(X = 170) \\
&= 0 \times 0.20 + 137 \times 0.55 + 170 \times 0.25 = 117.85
\end{aligned}
$$

Expected value of a Discrete Random Variable

If X takes outcomes x_1, ..., x_k with probabilities $P(X = x_1)$, ..., $P(X = x_k)$, the expected value of X is the sum of each outcome multiplied by its corresponding probability:

$$
\begin{aligned}
E(X) &= x_1 \times P(X = x_1) + \cdots + x_k \times P(X = x_k) \\
&= \sum_{i=1}^{k} x_i P(X = x_i) \qquad (2.71)
\end{aligned}
$$

The Greek letter μ may be used in place of the notation $E(X)$.



gure 2.23: A continuous distribution can also be balanced at its mean.

$$E(x) = \int_{-\infty}^{\infty} s\, f(s) ds$$

Expected Value

- Suppose X is a discrete random variable $P(X = a_i) = p_i$
  - The expected value of X is $E(X) = \sum_{i=1}^{n} p_i a_i$
- Suppose X is a continuous random variable with density f
  - The expected value of X is $E(X) = \int_{-\infty}^{+\infty} f(x) x dx$
- $E(X)$ is a property of the distribution, it is not a random variable.
- The average is a random variable:
  - $Average(x_1, x_2, \dots, x_n) \doteq \frac{1}{n}\sum_{i=1}^{n} x_i$
- When n is large, the average tends to be close to the mean.

Example - Binary random variables:

Let $X_1, X_2, \ldots, X_{100}$

Be independent binary random variables: $P(X_i = 0) = P(X_i = 1) = \frac{1}{2}$

Let $S = \frac{1}{100}\sum_{i=1}^{100} X_i$ S is the ________, S is/is-not a random variable?

$E(X_i) = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2}$, $E(X_i)$ is/is-not a random variable?

What is $E(S)$?

Rules for expected value:

1. If a,b are constants and X is a random variable then

$$E(aX+b)=aE(X)+b$$

2. If X,Y are random variables (dependent or independent)

$$E(X+Y)=E(X)+E(Y)$$

—> what is $E(aX+bY+c)=?$

3. If the distribution of the RV X is a mixture of two distributions:

$P=pP_1+(1-p)P_2$ then

$$E_P(X)=pE_{P_1}(X)+(1-p)E_{P_2}(X)$$

So now, $S=\frac{1}{100}\sum_{i=1}^{100}X_i$, what is $E(S)$?

Next Class

Expectation & Variance

CDFs Vs Histograms