RANDOM VARIABLES

Pau Fonseca, Ph. D. , Nihan Acar-Denizli, Ph.D.
Kristin L. Sainani, PhD

Random variables

Random variable

- We already know about **experiments**, **sample spaces**, and **events**. Now, we are interested in a number that is associated with the experiment. We conduct a random experiment E and after learning the outcome ω in S we calculate a number X.
  - That is, to each outcome ω in the sample space we associate a number $X(\omega) = x$.
- A random variable is a numeric description of the outcome of an experiment.
- A random variable X is a function $X : S \rightarrow R$ that associates to each outcome $\omega \in S$ exactly one number $X(\omega) = x$.

Random variables can be discrete or continuous

- Discrete random variables have a countable number of outcomes.
  - Examples: Dead/alive, treatment/placebo, dice, counts, the number of units sold, the number of costumers enter a bank by unit of time…
- Continuous random variables have an infinite continuum of possible values.
  - Examples: blood pressure, weight, the speed of a car, the real numbers from 1 to 6, time between arrivals…

Random Variable

- A variable X that takes “random” values. We assume that it follows a probability distribution, P(x).

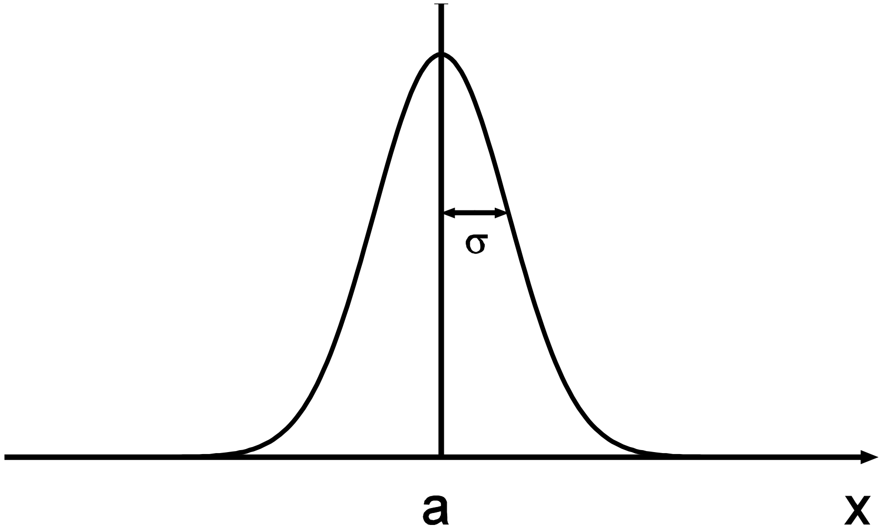
- Discrete variable: p_1, p_2, …

- Continuous variable: P(x)dx gives the probability that X falls between x and $x+dx$.

Probability functions

- A probability function, **maps** the possible values of x against their respective probabilities of occurrence, p(x).
  - p(x) is a number from 0 to 1.0.
  - The area under a probability function is always 1.

Gaussian (Normal) Distribution



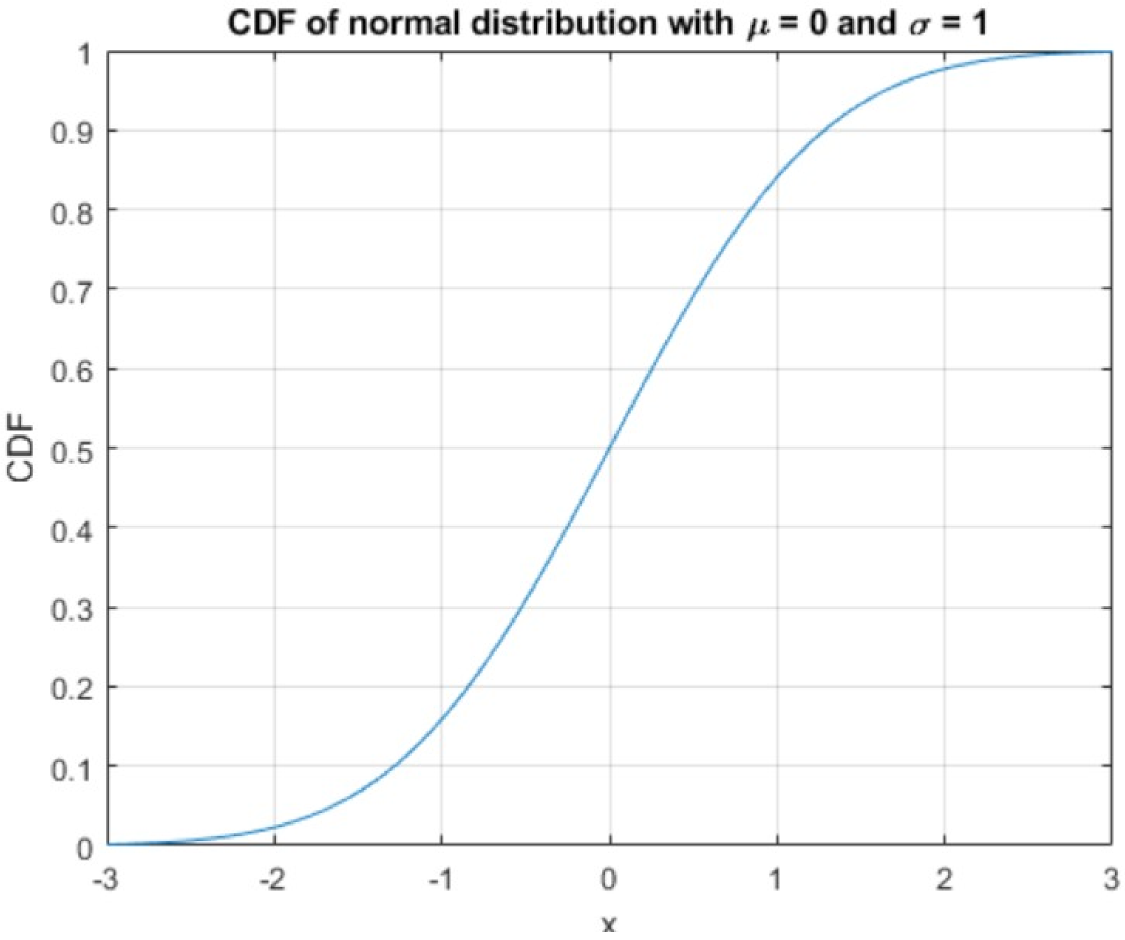
$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}}$$

Cumulative Distribution Function (CDF)

- The distribution function is defined as $F(x) = P(X \leq x)$.

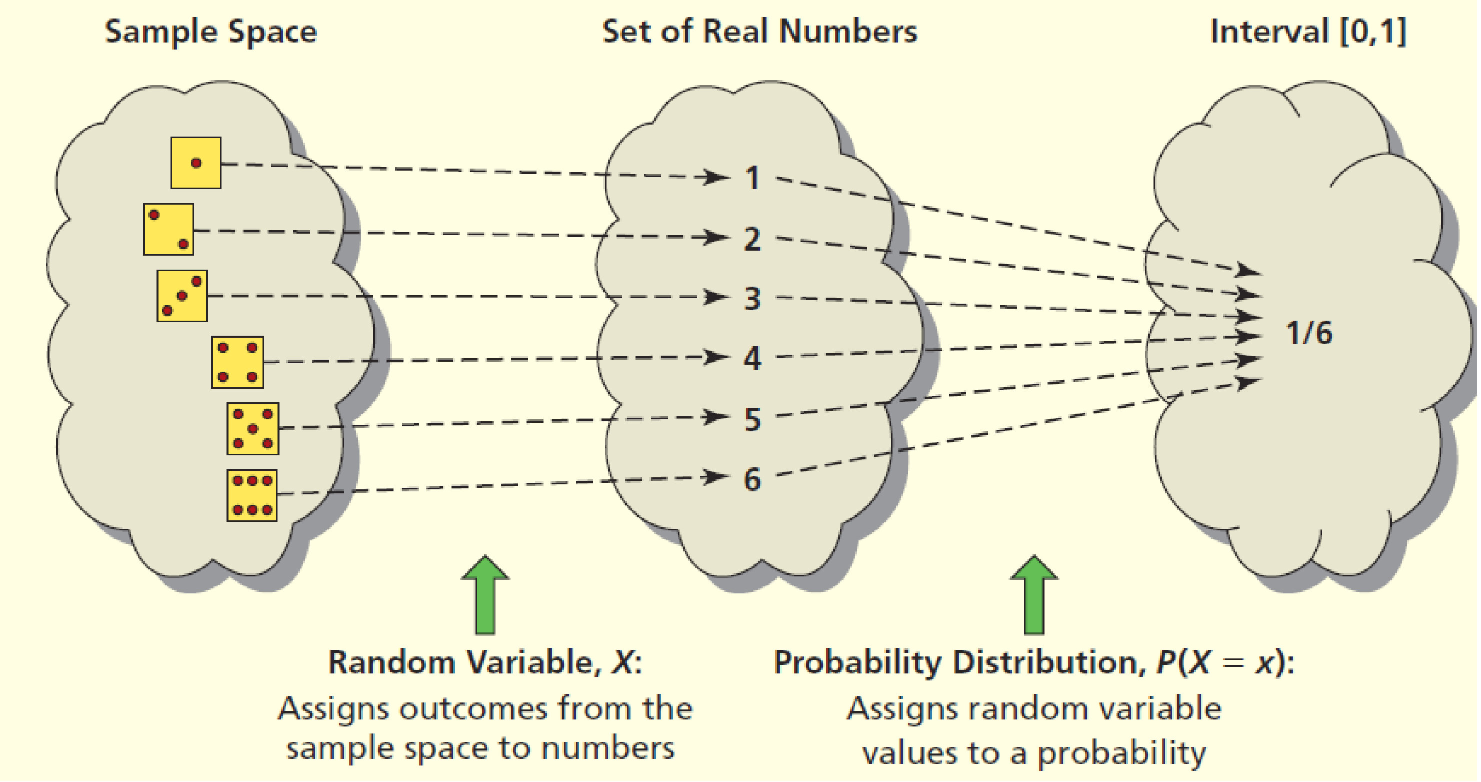
- This definition applies equally well for discrete and continuous random variables.

CDF of Standard Normal Distribution

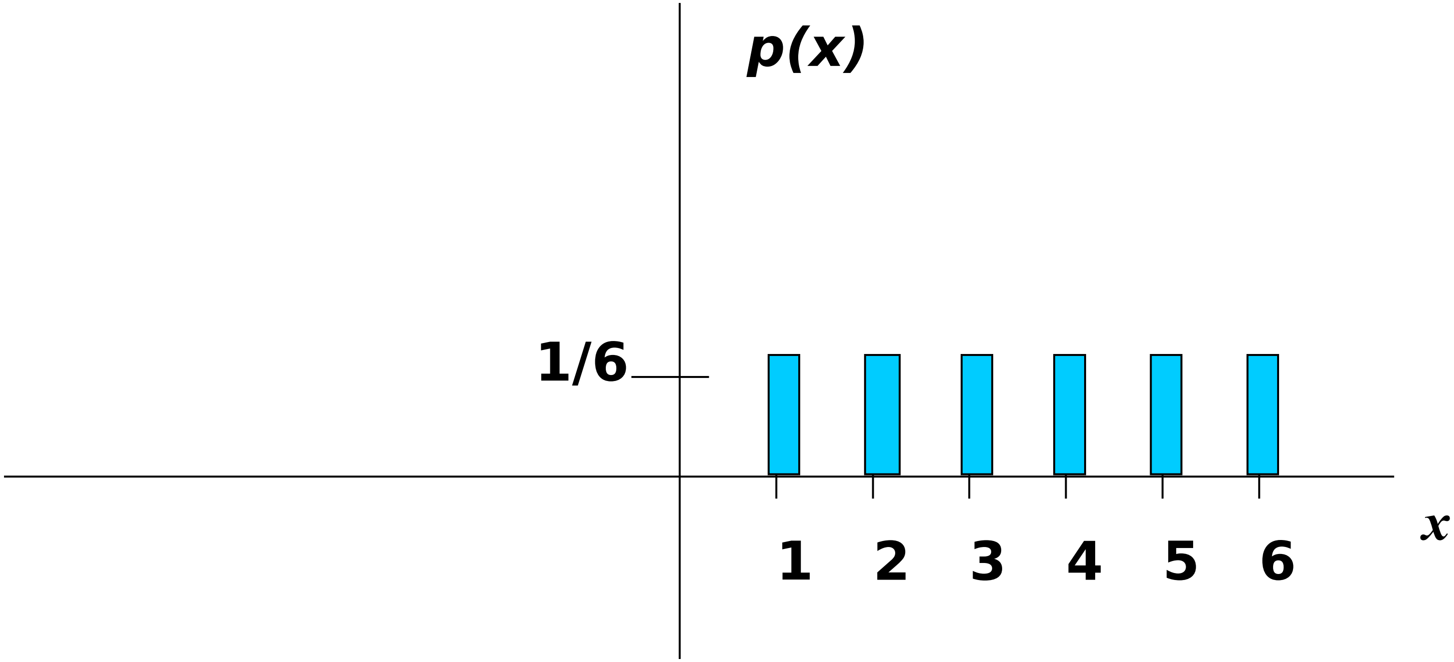


Discrete probability functions

Discrete Probability Function



Discrete example: roll of a die

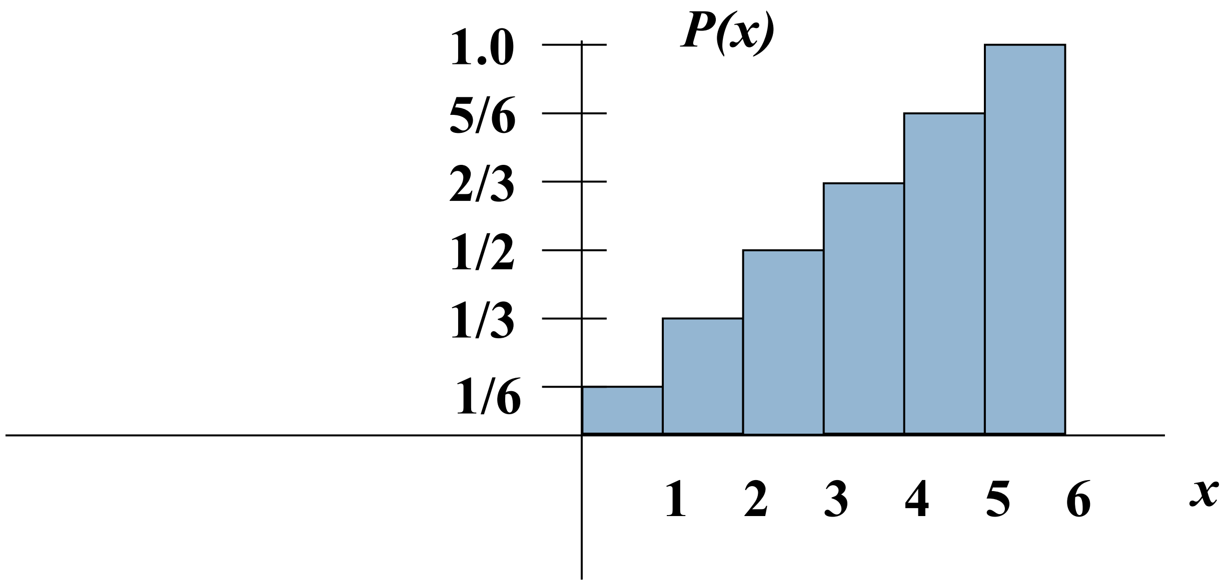


$$\sum_{\text{all } x} P(x) = 1$$

Probability mass function (pmf)

x	$p(x)$
1	$p(x=1) = 1/6$
2	$p(x=2) = 1/6$
3	$p(x=3) = 1/6$
4	$p(x=4) = 1/6$
5	$p(x=5) = 1/6$
6	$p(x=6) = 1/6$

Cumulative distribution function (CDF)



Cumulative distribution function

x	$P(x \leq A)$
1	$P(x \leq 1) = 1/6$
2	$P(x \leq 2) = 2/6$
3	$P(x \leq 3) = 3/6$
4	$P(x \leq 4) = 4/6$
5	$P(x \leq 5) = 5/6$
6	$P(x \leq 6) = 6/6$

Examples

- What's the probability that you roll a 3 or less?

  $P(x \leq 3) = 1/2.$

- What's the probability that you roll a 5 or higher?

  $P(x \geq 5) = 1 – P(x \leq 4) = 1\text{-}2/3 = 1/3.$

Practice Problem:

- The number of ships to arrive at a harbor on any given day is a random variable represented by x. The probability distribution for x is:

x	**10**	**11**	**12**	**13**	**14**
P(x)	**.4**	**.2**	**.2**	**.1**	**.1**

- Find the probability that on a given day:
  a) exactly 14 ships arrive: $p(x=14)=?$
  b) More than12 ships arrive: $p(x\geq 12)=?$
  c) Less than11 ships arrive: $p(x\leq 11)=?$

Practice Problem:

- The number of ships to arrive at a harbor on any given day is a random variable represented by x. The probability distribution for x is:

x	**10**	**11**	**12**	**13**	**14**
P(x)	**.4**	**.2**	**.2**	**.1**	**.1**

- Find the probability that on a given day:
  a) exactly 14 ships arrive: p(x=14)= 0.1
  b) More than12 ships arrive: p(x≥12)= (.2 + .1 +.1) = 0 .4
  c) Less than11 ships arrive: p(x≤11)= (.4 +.2) = 0 .6

Practice Problem:

- You are lecturing to a group of 1000 students. You ask them to each randomly pick an integer between 1 and 10. Assuming, their picks are truly random:
  - What’s your best guess for how many students picked the number 9?
  - What percentage of the students would you expect picked a number less than or equal to 6?

Practice Problem:

- You are lecturing to a group of 1000 students. You ask them to each randomly pick an integer between 1 and 10. Assuming, their picks are truly random:
  - What's your best guess for how many students picked the number 9?
    - Since $p(x=9) = 1/10$, we'd expect about 1/10th of the 1000 students to pick 9. Then, 100 students.
  - What percentage of the students would you expect picked a number less than or equal to 6?
    - Since $p(x \leq 6) = 1/10 + 1/10 + 1/10 + 1/10 + 1/10 + 1/10 = .6$ So, 60%.

Practice Problem

- Which of the following are probability functions?

Which of the following are probability functions?

a) $f(x) = 0.25$ for $x = 9,10,11,12$

a) for

b) $f(x) = \frac{3-x}{2}$ for $x = 1,2,3,4$

b) for

c) $f(x) = (x^2+x+1)/25$ for $x = 0,1,2,3$

c) for

Answer (a)

- a) $f(x) = 0.25$ for $x = 9,10,11,12$

x	*f(x)*
9	.25
10	.25
11	.25
12	.25

1.0

Yes, probability function!

Answer (b)

- b) $f(x) = \frac{3-x}{2}$ for $x = 1,2,3,4$

x	***f(x)***
1	**(3-1)/2=1.0**
2	**(3-2)/2=.5**
3	**(3-3)/2=0**
4	**(3-4)/2=-.5**

Though this sums to 1, you can't have a negative probability; therefore, it's not a probability function.

Answer (c)

- c) $f(x) = (x^2+x+1)/25$ for $x = 0,1,2,3$

x	f(x)
0	1/25
1	3/25
2	7/25
3	13/25

24/25

Doesn't sum to 1. Thus, it's not a probability function.

Important discrete distributions

- Binomial Distribution

  The probability of getting exactly *k* successes in *n* trials

  - Yes/no outcomes (dead/alive, treated/untreated, smoker/non-smoker, sick/well, etc.)

- Poisson Distribution

  The probability of x occurances in a given period.

  - Counts (e.g., how many cases of disease in a given area)

Binomial distribution

Probability Function of Binomial Distribution:

$$f(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}, \ x = 0, 1, ..., n$$

$n =$ number of trials

$p =$ probability of success on one trial

$x =$ number of successes in n trials

$f(x) =$ probability of x successes in n trials

Example: The probability of 4 customers making a purchase out of 10 customers given that the probability of a customer making a purchase is 0.30.

Poisson distribution

Probability Function of Poisson Distribution

$$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \ x = 1, 2, \ldots$$

λ = mean number of occurrences in an interval

$e = 2.71828$

x = number of occurrences in the interval

- Example: The probability that two customers will arrive to the bank per minute while we know that the mean number of customers arriving to the bank during corresponding period is 3.

Poisson Approximation to Binomial

- If n is "large" and p is "small" enough, then Poisson distribution can be used to approximate a Binomial distribution.
- Suppose 1000 women are screened for a rare type of cancer that has a nationwide incidence of 6 cases per 10,000 (i.e., $p = 0.0006$).
- What is the probability of finding two or fewer cases?
- Then, the Poisson mean is set equal to Binomial mean.

Poisson Approximation to Binomial

Poisson Approximation	*Actual Binomial Probability*
$P(X = 0) = 0.6^0 e^{-0.6}/0! = .5488$	$P(X = 0) = \dfrac{1000!}{0!(1000-0)!} .0006^0 (1 - .0006)^{1000-0} = .5487$
$P(X = 1) = 0.6^1 e^{-0.6}/1! = .3293$	$P(X = 1) = \dfrac{1000!}{1!(1000-1)!} .0006^1 (1 - .0006)^{1000-1} = .3294$
$P(X = 2) = 0.6^2 e^{-0.6}/2! = .0988$	$P(X = 2) = \dfrac{1000!}{2!(1000-2)!} .0006^2 (1 - .0006)^{1000-2} = .0988$

Poisson and Exponential

- When the arrival rates follow a Poisson process with mean arrival rate (λ) per unit of time, there will be on average x occurrences per t units of time.

- The Poisson distribution describing this process is:

$$P(x) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}$$

- The probability of no occurrences in t units of time:

$$P(x = 0) = e^{-\lambda t}$$

Poisson and Exponential

- It is possible then to interpret this as the probability that the time T, to the first occurrence is greater than t,

$$P(T > t) = P(x = 0|\mu = \lambda t) = e^{-\lambda t}$$

- Then, the probability that an event occur during t units of time is given by

$$P(T \leq t) = 1 - P(x = 0|\mu = \lambda t) = 1 - e^{-\lambda t}$$

- This is the exponential distribution with a mean time between arrivals of $1/\lambda$.

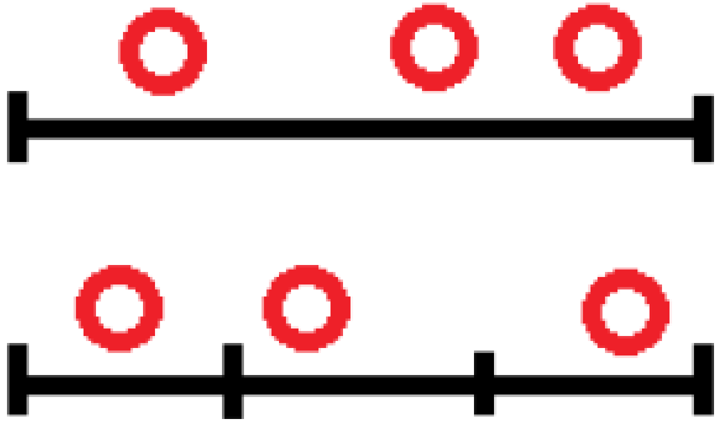
Poisson and Exponential

- Suppose that the number of cars arriving at a car wash during one hour is described by a Poisson distribution with a mean of 10 cars per hour.
- Since the average number of arrivals 10 cars per hour, the average time between cars is
- 1 hour/10 cars =0.1 hour/car

Model	*Random Variable*	*Parameter*	*Domain*	*Variable Type*
Poisson	$X =$ number of arrivals per unit of time	$\lambda = \frac{\text{(mean arrivals)}}{\text{(unit of time)}}$	$x = 0, 1, 2, \ldots$	Discrete
Exponential	$X =$ waiting time until next arrival	$\lambda = \frac{\text{(mean arrivals)}}{\text{(unit of time)}}$	$x \geq 0$	Continuous

Poisson and Exponential

- The exponential distribution, then, is also representative of a Poisson process but describes the time between arrivals and specifies that these time intervals are completely random.



Continuous probability functions

Continuous case

- The probability function that accompanies a continuous random variable is a continuous mathematical function that **integrates** to 1.
- The probabilities associated with continuous functions are just areas **under** the curve.
- Probabilities are given for a **range of values**, rather than a particular value

Continuous case

- For example, recall the negative exponential function (in probability, this is called an "exponential distribution"):

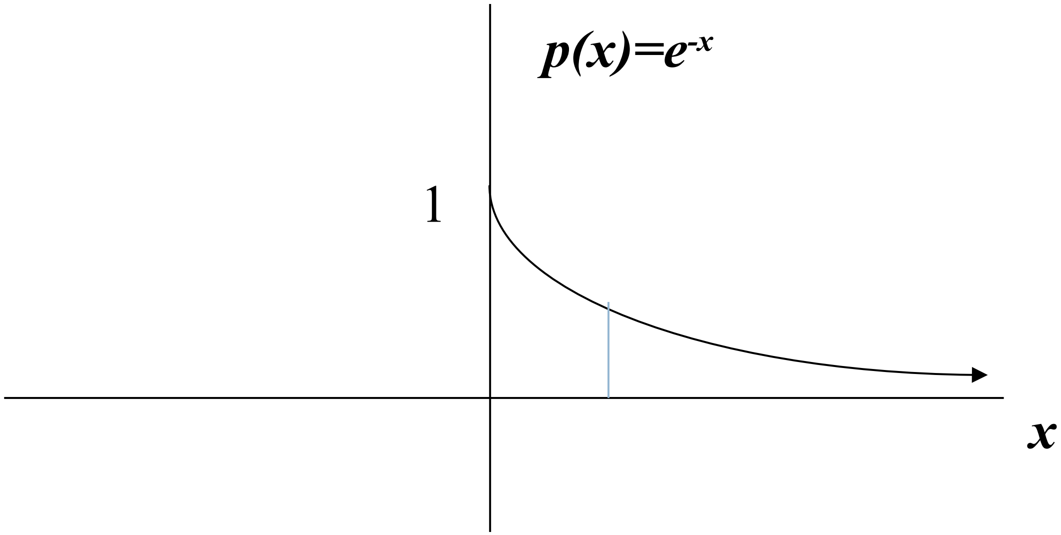
$$f(x) = e^{-x}$$

- This function integrates to 1

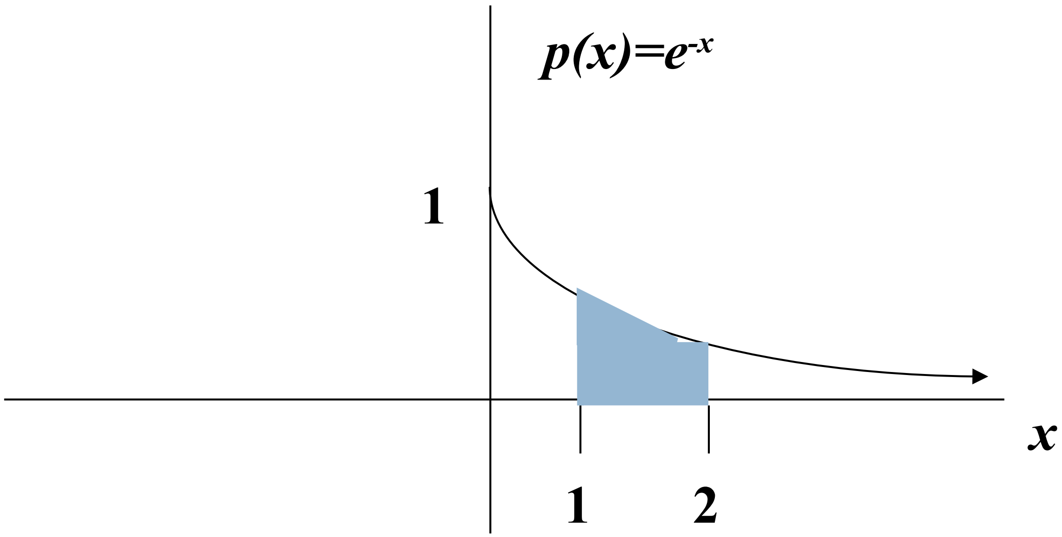
$$\int_0^{+\infty} e^{-x} = -e^{-x} \Big|_0^{+\infty} = 0 + 1 = 1$$

Continuous case: “probability density function” (pdf)

- The probability that x is any exact particular value (such as 1.9976) is 0; we can only assign probabilities to possible ranges of x.



For example, the probability of x falling within 1 to 2



$$P(1 \le x \le 2) = \int_1^2 e^{-x} = -e^{-x} \Big|_1^2 = -e^{-2} - -e^{-1} = -.135 + .368 = .23$$

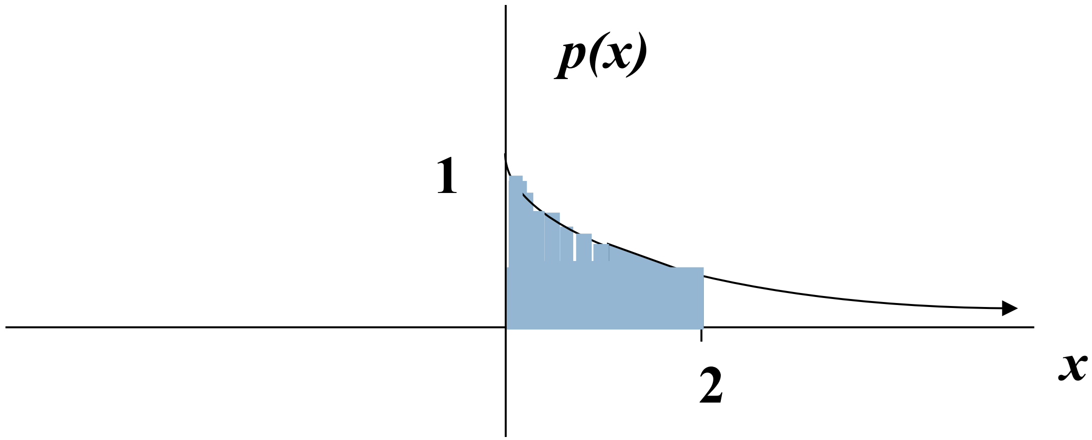
Cumulative distribution function

- As in the discrete case, we can specify the "cumulative distribution function" (CDF):

- The CDF here = P(x≤A)=

$$\int_0^A e^{-x} = -e^{-x} \Big|_0^A = -e^{-A} - -e^{0} = -e^{-A} + 1 = 1 - e^{-A}$$

Example

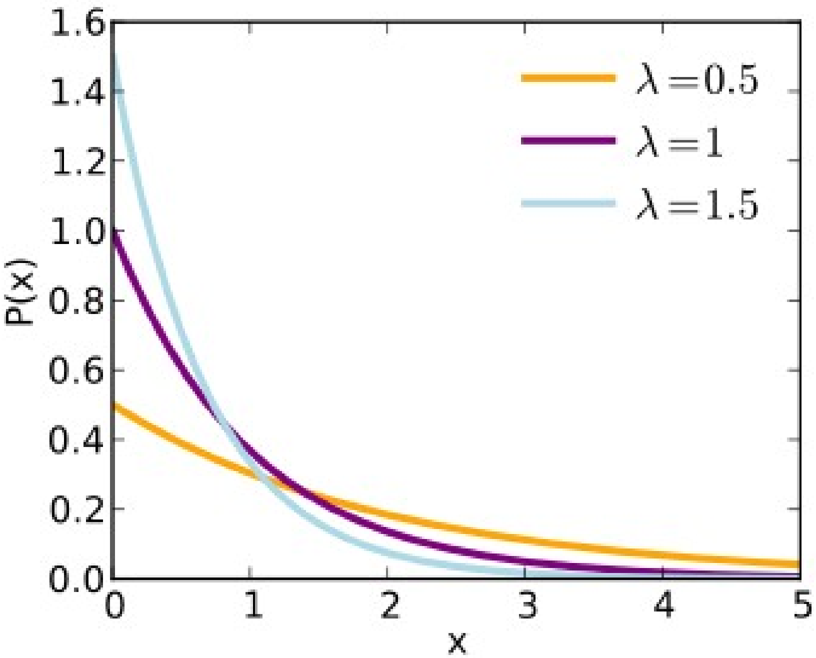


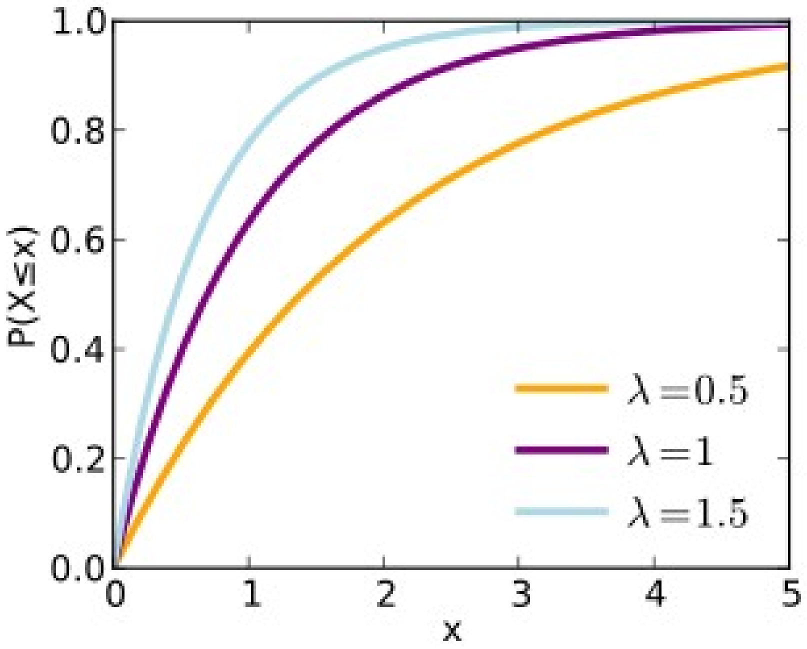
$$P(x \leq 2) = 1 - e^{-2} = 1 - .135 = .865$$

CDF and pdf for exponential

pdf

cdf



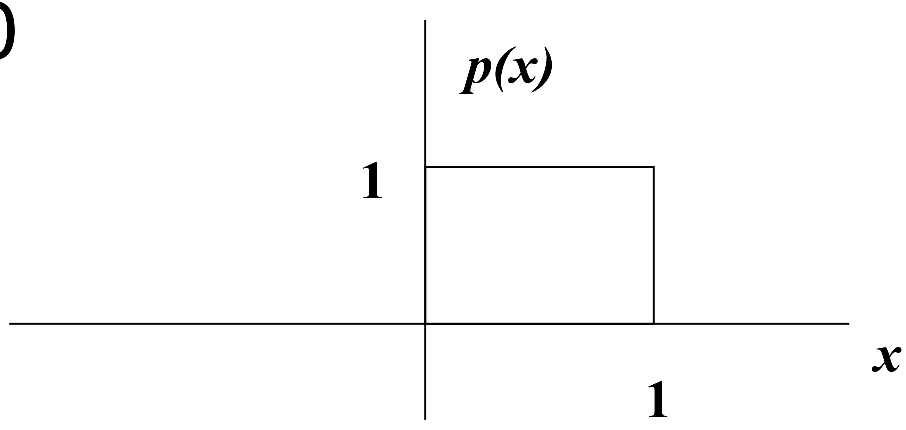


Example 2: Uniform distribution

- The uniform distribution: all values are equally likely.
- The pdf of Uniform distribution is:

  $f(x) = 1/b-a$, for $b \geq x \geq a$

- The uniform distribution: $f(x) = 1$, for $1 \geq x \geq 0$



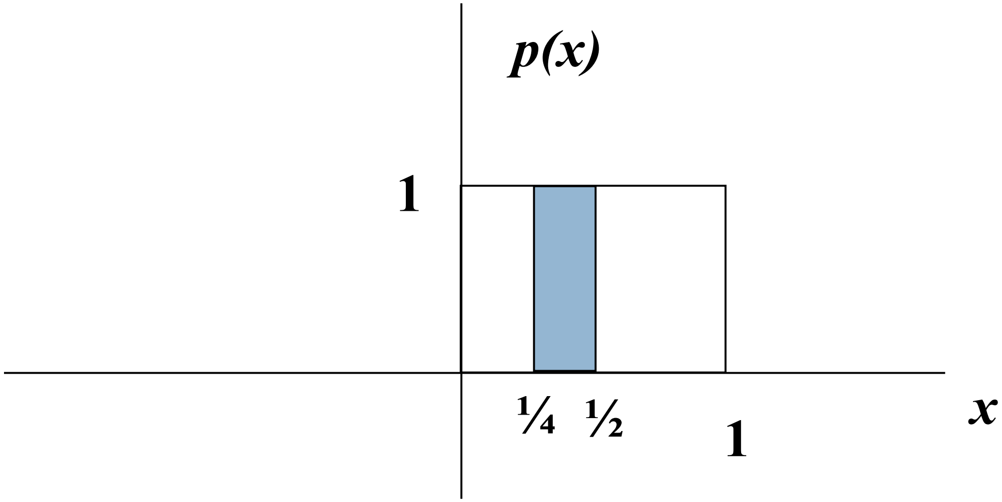
Example 2: Uniform distribution

- We can see it's a probability distribution because it integrates to 1 (the area under the curve is 1):

$$\int_0^1 1 = x \Big|_0^1 = 1 - 0 = 1$$

Example: Uniform distribution

- What's the probability that x is between ¼ and ½?



- P (½ ≥ x ≥ ¼) = ¼

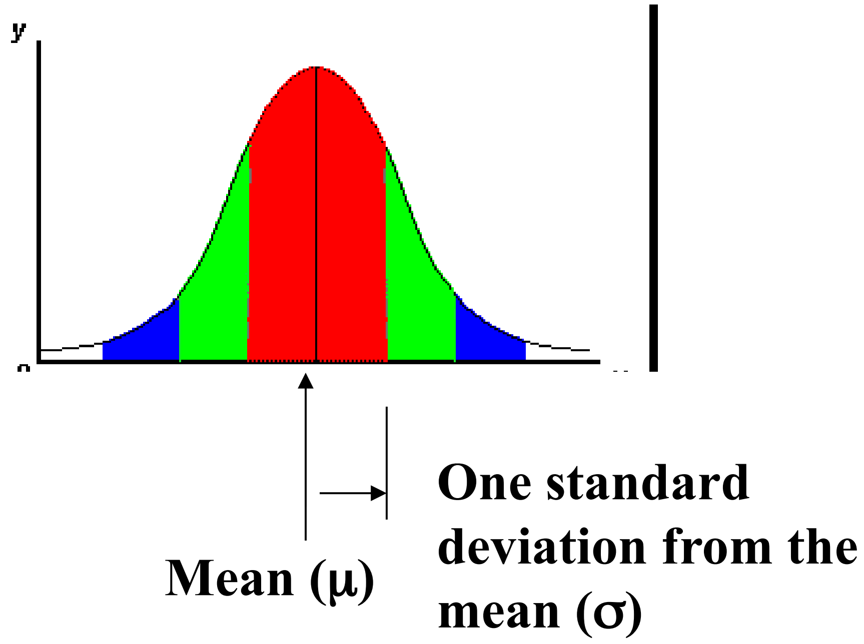
Expected value and variance

Statistics of a Random Variable

- Mean $<X> = (1/N) \sum x_i$
- Variance $\sigma^2 = <X^2> - <X>^2$
- Correlation $<X\,Y> - <X><Y$

Other higher moments are also useful:
- skewness
- kurtosis

For example, bell-curve (normal) distribution



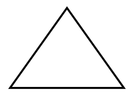
Expected value, or mean

- If we understand the underlying probability function of a certain phenomenon, then we can make informed decisions based on how we expect x **to behave on-average** over the long-run…(so called “frequentist” theory of probability).
- Expected value is just the **weighted average** or mean (μ) of random variable x. Imagine placing the masses p(x) at the points X on a beam; the balance point of the beam is the expected value of x.

Example: Expected value

- Recall the following probability distribution of ship arrivals:

x	**10**	**11**	**12**	**13**	**14**
P(x)	**.4**	**.2**	**.2**	**.1**	**.1**



$$\sum_{i=1}^{5} x_i\, p(x) = 10(.4) + 11(.2) + 12(.2) + 13(.1) + 14(.1) = 11.3$$

Empirical Mean is a special case of Expected Value...

- Sample mean, for a sample of n subjects: =

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \sum_{i=1}^{n} x_i \left(\frac{1}{n}\right)$$

The probability (frequency) of each person in the sample is 1/n.

Expected value, formally

- **Discrete case:**

$$E(X) = \sum_{\text{all x}} x_i p(x_i)$$

- **Continuous case:**

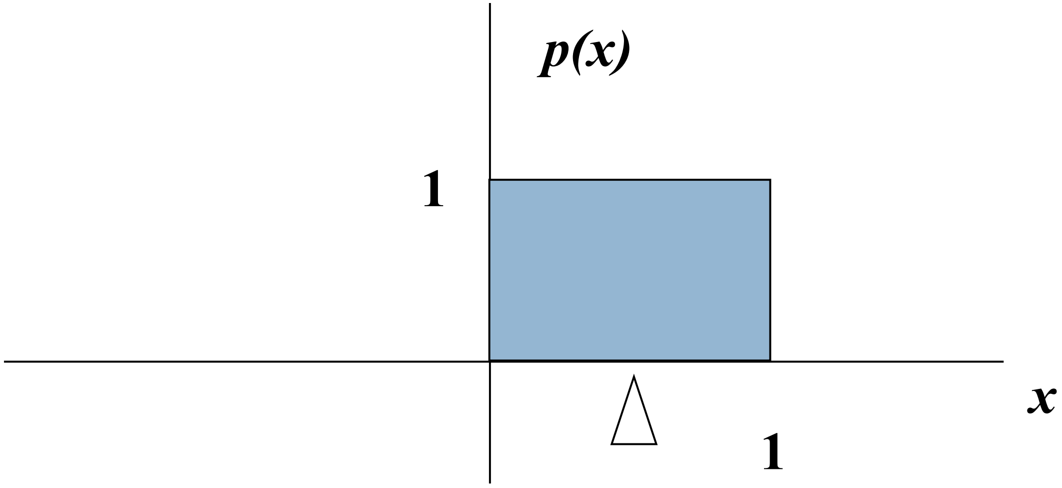
$$E(X) = \int_{\text{all x}} x_i p(x_i) dx$$

Expected Value

- If the probability distribution is known, the expected value (average value) can be computed as (for continuous variable)

$$\langle \mathrm{f}(X) \rangle = E\big(\mathrm{f}(X)\big) = \int \mathrm{f}(x)\,\mathrm{p}(x)\,dx$$

Extension to continuous case: uniform distribution

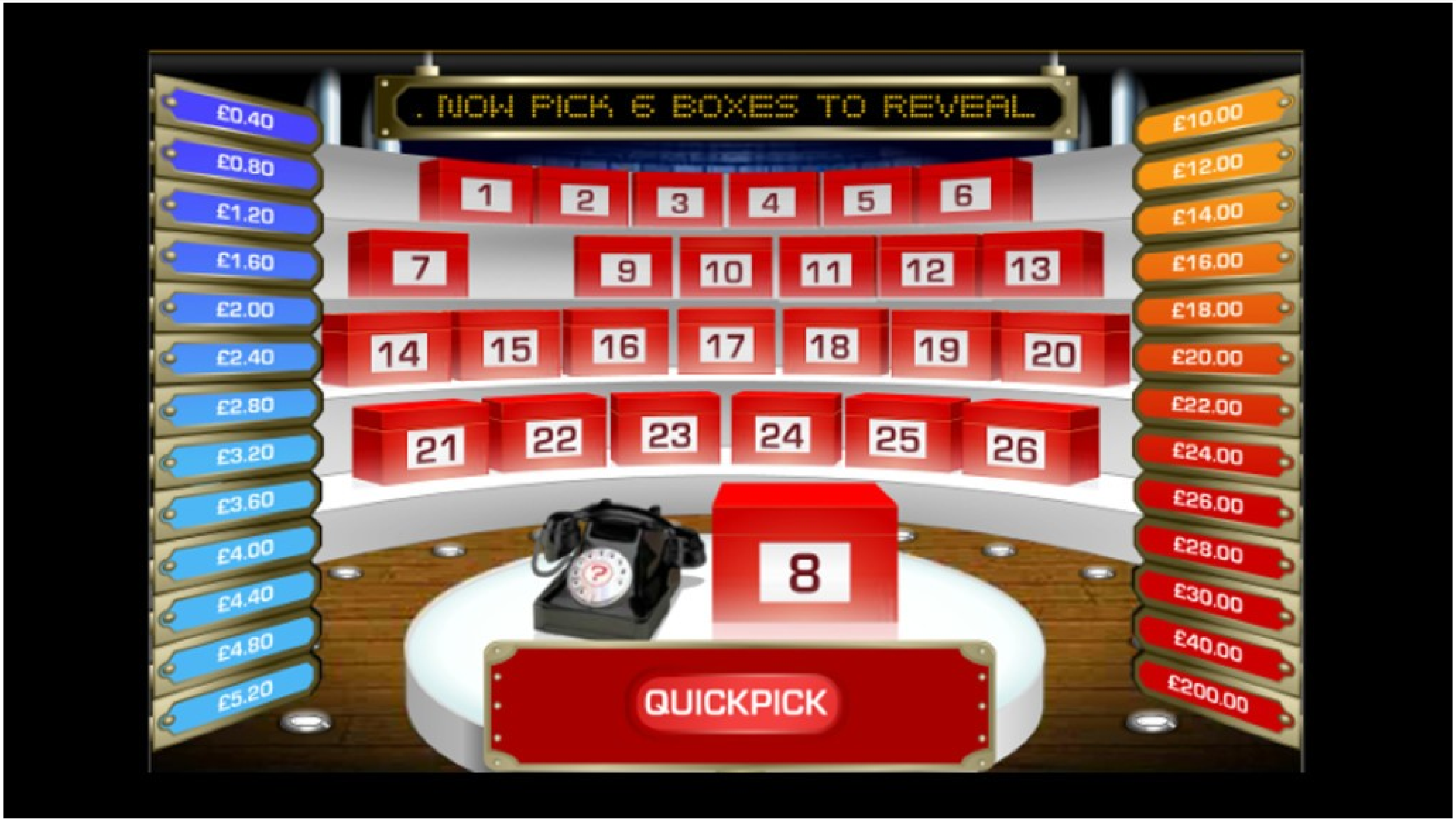


$$E(X) = \int_0^1 x(1)\,dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2} - 0 = \frac{1}{2}$$

Expected value

- $E(X) = \mu$
  - these symbols are used interchangeably.
  - Expected value is an extremely useful concept for good decision-making.

To Deal or Not to Deal?



Gambling

- A roulette wheel has the numbers 1 through 36, as well as 0 and 00. If you bet $1 that an odd number comes up, you win or lose $1 according to whether or not that event occurs.

  Let the random variable X denotes your net gain,

  X=1 with the probability18/38 and

  X= -1 with probability 20/38.

- If the cost is $10 per game, the casino wins an average of 53 cents per game.

  If 10,000 games are played in a night, that's a cool $5300.

Gambling

- A roulette wheel has the numbers 1 through 36, as well as 0 and 00. If you bet \$1 that an odd number comes up, you win or lose \$1 according to whether or not that event occurs. If random variable X denotes your net gain, X=1 with probability 18/38 and X= -1 with probability 20/38.
  - $E(X) = 1(18/38) - 1(20/38) = -\$.053$
  - On average, the casino wins (and the player loses) 5 cents per game.
- If the cost is \$10 per game, the casino wins an average of 53 cents per game. If 10,000 games are played in a night, that's a cool \$5300.
  - $E(X) = 10(18/38) - 10(20/38) = -\$.53$

Example: the lottery

- A certain lottery works by picking 6 numbers from 1 to 49. It costs $1.00 to play the lottery, and if you win, you win $2 million after taxes.

- If you play the lottery once, what are your expected winnings or losses?

Lottery

- Calculate the probability of winning in 1 try:

$$\frac{1}{\binom{49}{6}} = \frac{1}{\frac{49!}{43!6!}} = \frac{1}{13{,}983{,}816} = 7.2 \times 10^{-8}$$

"49 choose 6"

Out of 49 numbers, this is the number of distinct combinations of 6.

- The probability function (note, sums to 1.0):

x$	*p(x)*
-1	**.999999928**
+ 2 million	***7.2 x 10^--8^***

Expected Value

- The probability function

x$	***p(x)***
-1	**.999999928**
+ 2 million	***7.2 x 10^{--8}***

- Expected Value:
- E(X) = P(win)*$2,000,000 + P(lose)*-$1.00

  = (2.0×10^{6})* (7.2×10^{-8})+(0.999999928) (-1)

  = 0.144 - 0.999999928 = -$.86
- Negative expected value is never good!
- You shouldn’t play if you expect to lose money!

Expected Value

- If you play the lottery every week for 10 years, what are your expected winnings or losses?

  - 520 x (-.86) = -$447.20

Expected Value as a mathematical operator

- If c= a constant number (i.e., not a variable) and X and Y are any random variables…
- $E(c) = c$
  - No randomness. You always expect to get c.
- $E(cX) = c\,E(X)$
- $E(c + X)=c + E(X)$
- $E(X+Y)= E(X) + E(Y)$

$E(cX)=c\,E(X)$

- $E(cX)=c\,E(X)$
- Example: If the casino charges $10 per game instead of $1, then the casino expects to make 10 times as much on average from the game.

$E(c + X) = c + E(X)$

- $E(c + X) = c + E(X)$
- Example, if the casino throws in a free drink worth exactly \$5.00 every time you play a game, you always expect to (and do) gain an extra \$5.00 regardless of the outcome of the game.

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

- $E(X+Y)= E(X) + E(Y)$
- Example: If you play the lottery twice, you expect to lose: -\$.86 + -\$.86.
  - This works even if X and Y are dependent.
  - Does not require independence.

Variance/standard deviation

- "The average (expected) squared distance (or deviation) from the mean"

$$\sigma^2 = Var(x) = E[(x - \mu)^2] = \sum_{\text{all x}} (x_i - \mu)^2 p(x_i)$$

Variance, formally

- **Discrete case:**

$$Var(X) = \sigma^2 = \sum_{\text{all x}} (x_i - \mu)^2 p(x_i)$$

- **Continuous case:**

$$Var(X) = \sigma^2 = \int_{-\infty}^{\infty} (x_i - \mu)^2 p(x_i) dx$$

Similarity to empirical variance

- The variance of a sample:

The variance of a sample:

$$S^2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{n - 1} = \sum_{i=1}^{N}(x_i - \bar{x})^2 \left(\frac{1}{n - 1}\right)$$

Division by n-1 reflects the fact that we have lost a “degree of freedom” (piece of information) because we had to estimate the sample mean before we could estimate the sample variance.

- $Var(X) = \sigma^2$
  - these symbols are used interchangeably

these symbols are used interchangeably

Practice Problem

- A roulette wheel has the numbers 1 through 36, as well as 0 and 00. If you bet \$1.00 that an odd number comes up, you win or lose \$1.00 according to whether or not that event occurs.
- If X denotes your net gain, X=1 with probability 18/38 and X= -1 with probability 20/38.
  - We already calculated the mean to be = -\$.053. What's the variance of X?

Answer

- Standard deviation is \$.99. Interpretation: On average, you're either 1 dollar above or 1 dollar below the mean, which is just under zero. Makes sense!

$$\sigma^2 = \sum_{\text{all x}} (x_i - \mu)^2 p(x_i)$$

$$= (+1 - -.053)^2 (18/38) + (-1 - -.053)^2 (20/38)$$

$$= (1.053)^2 (18/38) + (-1 + .053)^2 (20/38)$$

$$= (1.053)^2 (18/38) + (-.947)^2 (20/38)$$

$$= .997$$

$$\sigma = \sqrt{.997} = .99$$

Handy calculation formula

$$Var(X) = \sum_{\text{all x}} (x_i - \mu)^2 p(x_i) = \sum_{\text{all x}} x_i^2 p(x_i) - (\mu)^2$$

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

Handy calculation formula

- $E(x-\mu)^2 = E(x^2 - 2\mu x + \mu^2)$

  $= E(x^2) - E(2\mu x) + E(\mu^2)$

- Use rules of expected value: E(X+Y) = E(X) + E(Y)
  - $= E(x^2) - 2\mu E(x) + \mu^2$ // $(E(c) = c)$
  - $= E(x^2) - 2\mu\mu + \mu^2$ // $(E(x) = \mu)$
  - $= E(x^2) - \mu^2$
  - $E(x^2) - [E(x)]^2$

Handy calculation formula

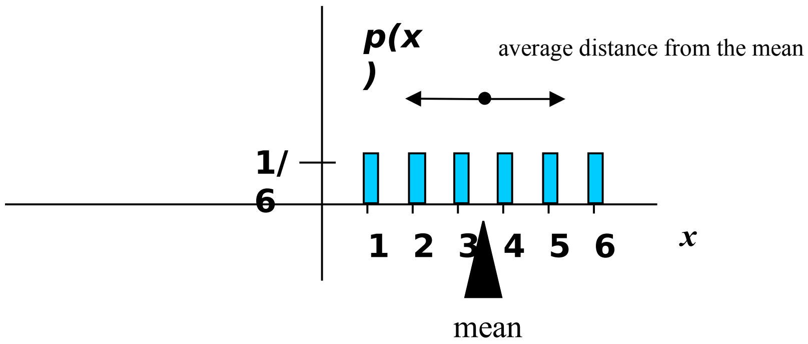
- OR, equivalently:

- OR, equivalently:

- $$
\begin{aligned}
Var(x) &= \sum[(x-\mu)^2]p(x) \\
&= \sum[x^2 - 2\mu x + \mu^2]p(x) \\
&= \sum x^2 p(x) - 2\mu \sum x p(x) + \mu^2 \sum p(x) \\
&= E(x^2) - 2\mu E(x) + \mu^2(1) \\
&= E(x^2) - 2\mu^2 + \mu^2 \\
&= \mathrm{E}(\mathrm{x}^2) - \mu^2 \\
&= E(x^2) - [E(x)^2]
\end{aligned}
$$

For example, what's the variance and standard deviation of the roll of a die?

x	p(x)
1	p(x=1)=1/6
2	p(x=2)=1/6
3	p(x=3)=1/6
4	p(x=4)=1/6
5	p(x=5)=1/6
6	p(x=6)=1/6
	1.0



$$E(x) = \sum_{\text{all x}} x_i\, p(x_i) = (1)(\frac{1}{6}) + 2(\frac{1}{6}) + 3(\frac{1}{6}) + 4(\frac{1}{6}) + 5(\frac{1}{6}) + 6(\frac{1}{6}) = \frac{21}{6} = 3.5$$

$$E(x^2) = \sum_{\text{all x}} x_i^{\,2}\, p(x_i) = (1)(\frac{1}{6}) + 4(\frac{1}{6}) + 9(\frac{1}{6}) + 16(\frac{1}{6}) + 25(\frac{1}{6}) + 36(\frac{1}{6}) = 15.17$$

$$\sigma_x^2 = Var(x) = E(x^2) - [E(x)]^2 = 15.17 - 3.5^2 = 2.92$$

$$\sigma_x = \sqrt{2.92} = 1.71$$

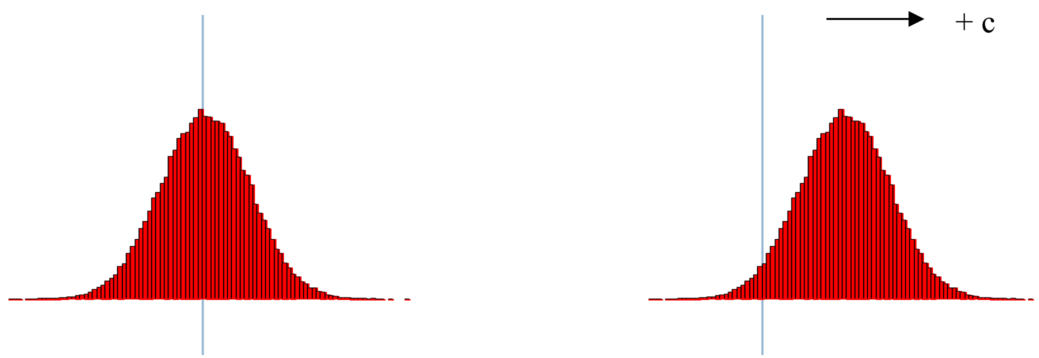
Variance as a mathematical operator:

- If c = a constant number (i.e., not a variable) and X and Y are random variables, then
- Var(c) = 0
- Var (c+X)= Var(X)
- Var(cX)= c^2Var(X)
- Var(X+Y)= Var(X) + Var(Y)
  ONLY IF X and Y are independent
- Var(X+Y)= Var(X) + Var(Y)+2Cov(X,Y)
  IF X and Y are not independent}

Var $(c + X) = \text{Var}(X)$

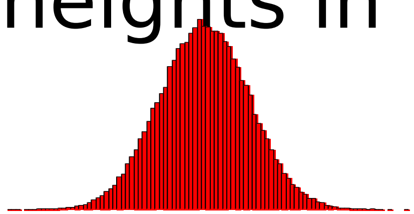
- Var $(c + X) = \text{Var}(X)$
- Adding a constant to every instance of a random variable doesn't change the variability. It just shifts the whole distribution by c. If everybody grew 5 inches suddenly, the variability in the population would still be the same.

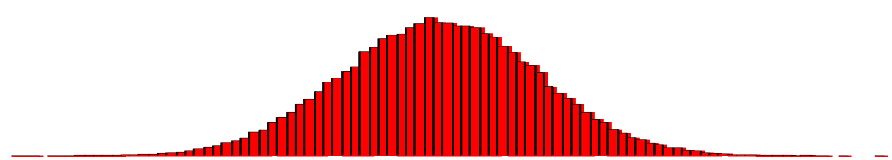
- Var $(c + X) = \text{Var}(X)$
- Adding a constant to every instance of a random variable doesn't change the variability. It just shifts the whole distribution by c. If everybody grew 5 inches suddenly, the variability in the population would still be the same.



$Var(cX) = c^2 Var(X)$

- $Var(cX) = c^2 Var(X)$
- Multiplying each instance of the random variable by c makes it c-times as wide of a distribution, which corresponds to c^2 as much variance (deviation squared).
- For example, if everyone suddenly became twice as tall, there'd be twice the deviation and 4 times the variance in heights in the population.





$Var(X+Y)= Var(X) + Var(Y)$

- $Var(X+Y)= Var(X) + Var(Y)$

  ONLY IF X and Y are independent!!!!!!!!

- With two random variables, you have more opportunity for variation, unless they vary together (are dependent, or have covariance):

  $Var(X+Y)= Var(X) + Var(Y) + 2Cov(X, Y)$

To know more

- Chapter 3: Discrete Random Variables and Their Distributions and Chapter 4: Continuous Distributions of **Variables of Probability and Statistics and for Computer Scientists** (2014 Ed.)