More on Cumulative Probability Function (CDF)

General Info

Reminder: RVs

Discrete?

Terms & Notation

More on CDF

> The *cumulative probability distribution (cdf)* for a random variable X is a function $F(x)$ that for each number x gives the probability that X takes that value or a smaller one,
> $F(x) = P[X \leq x]$.

Since (for discrete distributions) probabilities are calculated by summing values of $f(x)$,

$$F(x) = P[X \leq x] = \sum_{y \leq x} f(y)$$

General Info

Reminder: RVs

Discrete?

Terms & Notation

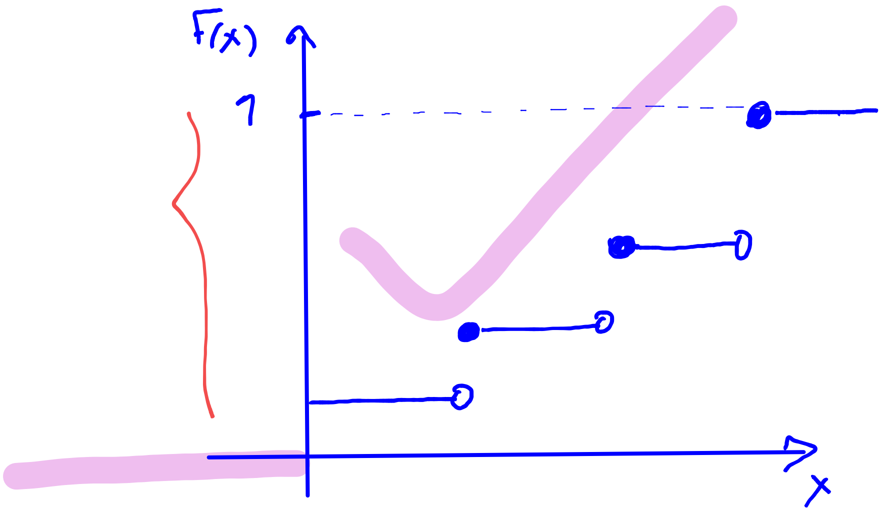
More on CDF

Properties of a mathematically valid cumulative distribution function:

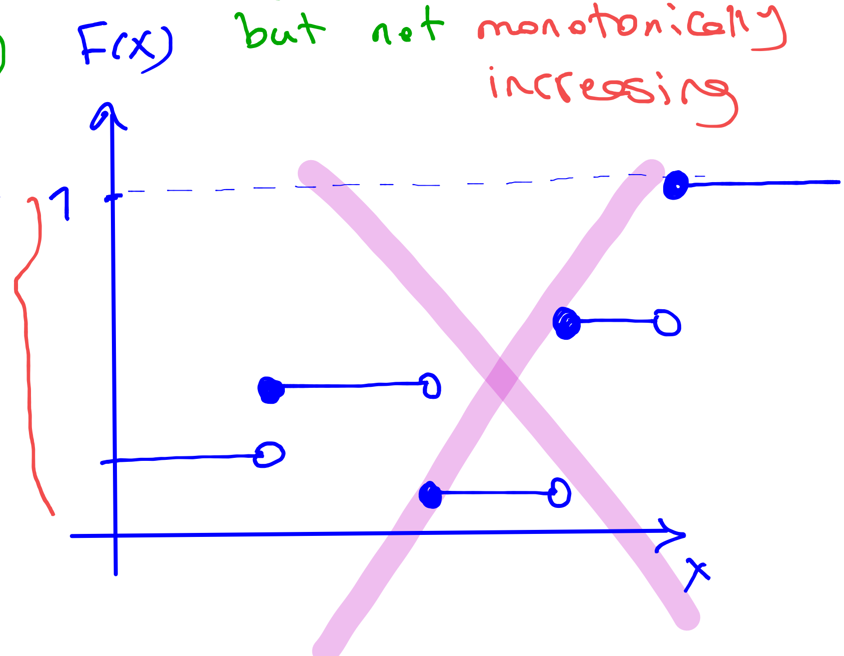
- $F(x) \geq 0$ for all real numbers x
- $F(x)$ is monotonically **increasing**
- $F(x)$ is right continuous
- $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to +\infty} F(x) = 1$ $\equiv$ $0 \leq F(x) \leq 1$
  - This means that $0 \leq F(x) \leq 1$ for **any CDF**

> In the discrete cases, the graph of $F(x)$ will be n stair-step graph with jumps at possible values of our random variable and height equal to the probabilities associated with those values

Right continuous & monotonically increasing



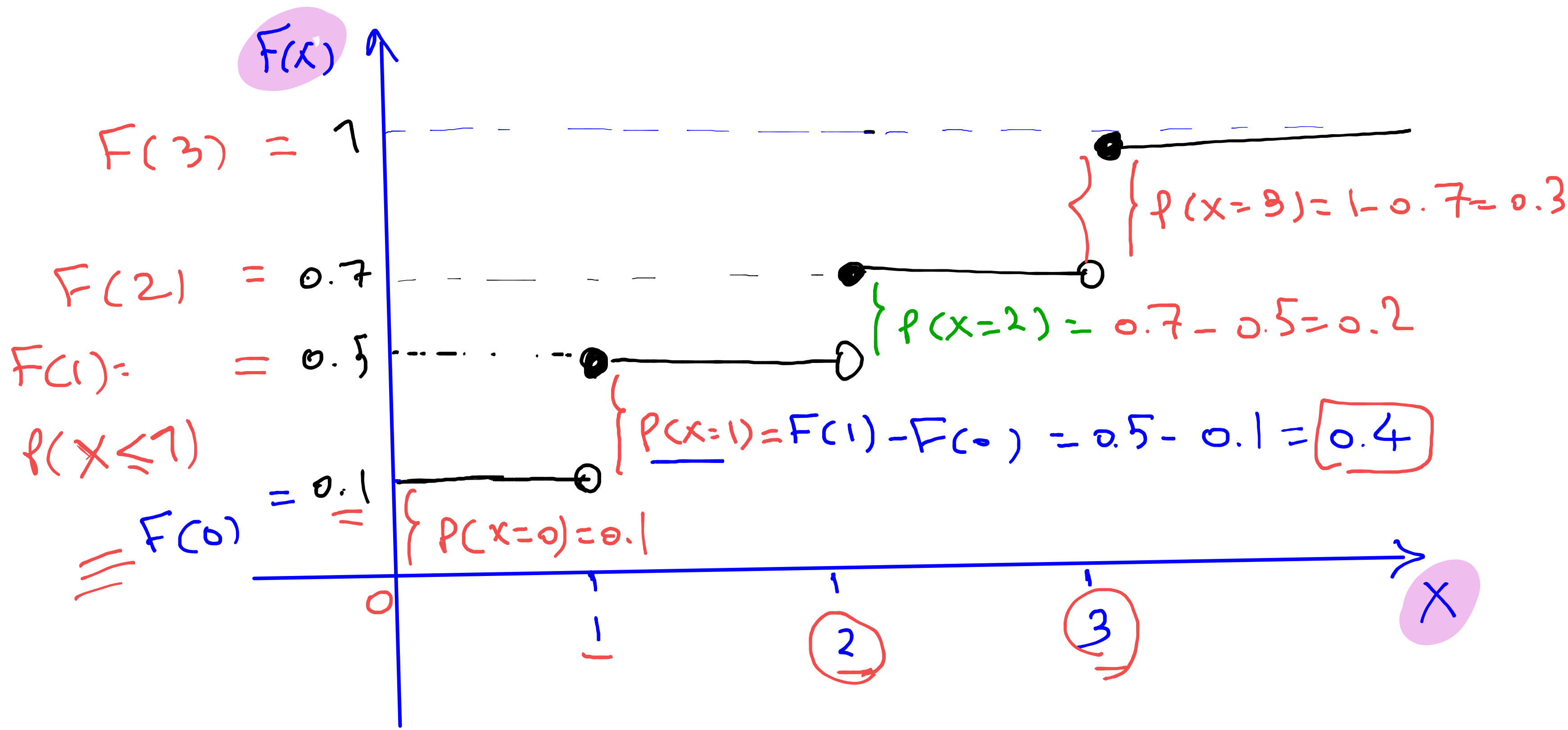
Right continuous, but not monotonically increasing



always: $0 \le F(x) \le 1$

Note: The height of each step (jump) is the probability $P(X=x)$ in

Discrete CDF:



$P(X \le 0)$

$$P(X=x) = f(x) = \begin{cases} 0.1 & x=0 \\ 0.4 & x=1 \\ 0.2 & x=2 \\ 0.3 & x=3 \end{cases}$$

$$= F(x) = \begin{cases} 0.1 & x=0 \\ 0.5 & x=1 \\ 0.7 & x=2 \\ 1 & x=3 \end{cases}$$

$$P(X \le 1) = P(X=0) + P(X=1)$$

General Info

Reminder: RVs

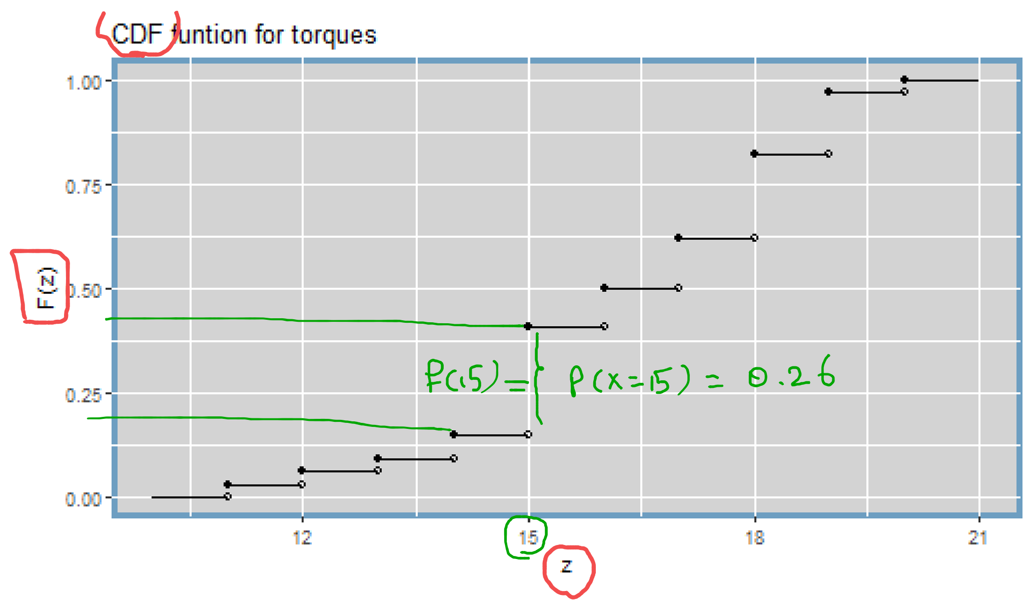
Discrete? $P(Z=z) =$

Terms & Notation

More on CDF

Example: [Torque] Let $Z =$ the torque, rounded to the nearest integer, required to loosen the next bolt on an apparatus.

Z	11	12	13	14	15	16	17	18	19	20
f(z)	0.03	0.03	0.03	0.06	0.26	0.09	0.12	0.20	0.15	0.03



General Info

Reminder: RVs

Discrete?

Terms & Notation

More on CDF

Calculate the following probabilities using the **cdf only**:

- $F(10.7)$ $= P(Z \le 10.7) = 0$

- $P(Z \le 15.5)$ $= F(15.5) = F(15) = \sum_{z=11}^{15} f(z)$ $\quad P(Z=z)$

  $= 0.41$

- $P(12.1 < Z \le 14)$ $= P(13 \le Z \le 14)$

  $= P(Z=13 \text{ or } Z=14) = P(Z=3) + P(Z=14)$

  $= f(13) + f(14) = 0.03 + 0.06$

  $= 0.09$

- $P(15 \le Z < 18)$

  $= P(15 \le Z \le 17) = P(Z=15 \text{ or } Z=16 \text{ or } Z=17)$

  $= f(15) + f(16) + f(17) = 0.47$

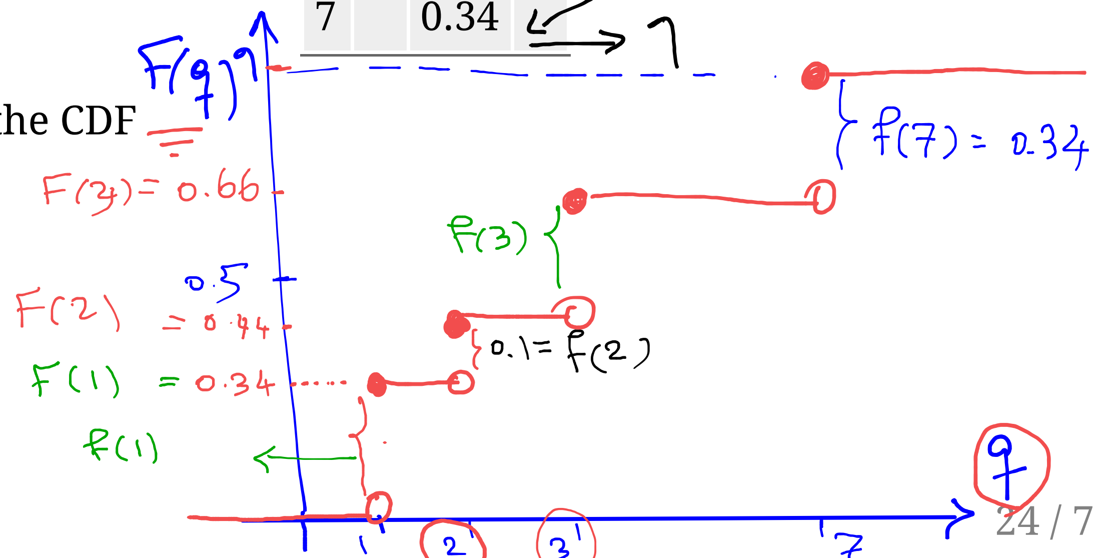
More on CDF

One more example

Say we have a random variable Q with pmf:

q		f(q)		F(q)
1		0.34		0.34
2		0.10		0.44
3		0.22		0.66
7		0.34		1

Draw the CDF



General Info

Reminder: RVs

Discrete?

Terms & Notation

Summaries

Almost all of the devices for describing relative frequency (empirical) distributions in Ch. 3 have versions that can describe (theoretical) probability distributions.

1. Measures of location == Mean
2. Measures of spread == variance
3. Histogram == probability histograms based on theoretical probabilities

Mean

and

Variance

of Discrete Random Variables

General Info

Reminder: RVs

Discrete?

Terms & Notation

Mean of a Discrete Random Variable

For a discrete random varable, X, which can take values $x_1, x_2, \ldots$ we define **the mean of X** (also known as **the expected value of X**) as:

$$E(X) = \sum_{i=1}^{n} x_i \cdot f(x_i)$$

We often use the symbol μ instead of $E(X)$.

Also, just to be confusing, you will often see EX instead of $E(X)$. Use context clues.

Example:

Suppose that the we roll a die and let T be the number of dots facing up. Find the expected value of T.

$$E(T) = \sum_{t=1}^{6} t \underline{f(t)} = (1)(1/6) + (2)(1/6) + (3)(1/6) + \cdots + (6)(1/6)$$

$$= 3.5$$

An intuitive way to get Expected value and variance of any random variable:

For discrete r.v: say we have a pmf $f(x)$:

$$E(\text{"anything"}) = \sum_{\forall x \in S} (\text{value of "anything"}) \cdot f(x)$$

e.g:

$$E(X) = \sum_{x \in S} x f(x) = \sum_{x \in S} x \cdot P(X = x)$$

$$E(X^2) = \sum_{x \in S} x^2 f(x) = \sum_{x \in S} x^2 P(X = x)$$

$$E(\sqrt{X}) = \sum_{x \in S} \sqrt{x}\, f(x) \equiv \sum_{x \in S} \sqrt{x}\, P(X = x)$$

e.g. For rolling a die:

T	1	2	3	4	5	6
$P(T=t) \equiv f(t)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

$$E(T) = \sum_{t \in \{1,2,\cdots,6\}} t\, f(t) = \sum_{\in \{1,2,3\cdots,6\}} t\, f(t) = (1)f(1) + (2)f(2) + (3)f(3) + (4)f(4) + (5)f(5) + (6)f(6)$$

$$= (1)(1/6) + (2)(1/6) + (3)(1/6) + (4)(1/6) + (5)(1/6) + (6)(1/6)$$

(Facto out $1/6$) $$= 1/6\,(1+2+3+4+5+6) = 21/6 = \underline{3.5}$$

$$E(\sqrt{T}) = \sum_{t \in \{1,2,\dots,6\}} \sqrt{t}\, f(t) = (\sqrt{1}) f(1) + \sqrt{2} f(2) + \sqrt{3} f(3)$$

$$+ \sqrt{4} f(4 + \sqrt{5} f(5) + \sqrt{6} f(6)$$

$$= \sqrt{1}(1/6) + \sqrt{2}(1/6) + \sqrt{3}(1/6) + \sqrt{4}(1/6) + \sqrt{5}(1/6) + \sqrt{6}(1/6)$$

$$= 1/6(\sqrt{1} + \sqrt{2} + \sqrt{3} + \sqrt{4} + \sqrt{5} + \sqrt{6})$$

$$E(T^2) = \sum_{t \in \{1,2,\dots,6\}} t^2 f(t) = 1^2 f(1) + 2^2 f(2) + 3^2 f(3) + 4^2 f(4) + 5^2 f(5)$$

$$+ 6^2 f(6)$$

$$= 1 \cdot 1/6 + 4 \cdot 1/6 + 9 \cdot 1/6 + 16 \cdot 1/6 + 25 \cdot 1/6 + 36 \cdot 1/6$$

$$= 1/6(1+4+9+16+25+36) = 91/6 = 15.166$$

Variance of a Discrete Random Variable

For a discrete random varable, X, which can take values $x_1, x_2, \ldots$ and has mean μ we define **the variance of X** as:

$$E[(X-\mu)^2] = Var(X) = \sum_{i=1}^{n} (x_i - \underset{E(X)}{\mu})^2 \cdot f(x_i)$$

There are other usefule ways to write this, most importantly:

$$Var(X) = \sum_{i=1}^{n} x_i^2 \cdot f(x_i) - \underbrace{\mu^2}_{[E(X)]^2}$$

which is the same as

$$\mathrm{Var} X = \sum_{x} (x - \mathrm{E}X)^2 f(x) = \mathrm{E}(X^2) - (\mathrm{E}X)^2.$$

$$E(T^2) = \sum_{\forall t \in S} t^2 f(t)$$

General Info

Reminder: RVs

Discrete?

Terms & Notation

Variance of a Discrete Random Variable

Example:

Suppose that the we roll a die and let T be the number of dots facing up. What is the variance of T?

T	1	2	3	4	5	6
$f(t)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$
$t\,f(t)$	$1/6$	$2/6$	$3/6$	$4/6$	$5/6$	$6/6$
$t^2 f(t)$	$1/6$	$4/6$	$9/6$	$16/6$	$25/6$	$36/6$

$$E(T) = \sum_{t \in \{1,\dots,6\}} t\,f(t) = 21/6 \quad , \quad E(T^2) = \sum_{t \in \{1,\dots,6\}} t^2 f(t) = 91/6$$

So, using the Formula of variance:

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

$$= 91/6 - [21/6]^2 = 2.91$$

Do Not make big mistakes:

① for any r.v (X): $0 \leq P(X=x) \leq 1$ (always)

② for any r.v (X): $0 \leq F(x) \leq 1$ (CDF)

③ Unlike probabilities & CDF, $E(X)$ can take any value $\in (-\infty, +\infty)$

But $var(X) \geq 0$ (always) i.e $var(X) \in (0, +\infty)$

General Info

Reminder: RVs

Discrete?

Terms & Notation

Variance of a Discrete Random Variable

Example

Say we have a random variable Q with pmf:

q	**f(q)**
1	0.34
2	0.10
3	0.22
7	0.34

Find the variance and standard deviation $= \sqrt{var(Q)}$

$$var(Q) = E(Q^2) - [E(Q)]^2$$

So, we need to find $E(Q)$ & $E(Q^2)$

by formulas,

$$E(Q) = \sum_{q \in \{1,2,3,7\}} q f(q) = (1)f(1) + (2)f(2) + (3)f(3) + (7)f(7)$$

$$= (1)(.34) + (2)(.1) + (3)(.22) + (7)(.34)$$

$$= .34 + .2 + .66 + 2.38$$

$$= 3.58$$

$$E(Q^2) = \sum_{q \in \{1,2,3,7\}} q^2 f(q) = (1)^2 f(1) + (2)^2 f(2) + (3)^2 f(3) + (7)^2 f(7)$$

$$= (1)(0.34) + (4)(0.1) + (9)(0.22) + 49(0.34)$$

$$= 0.34 + 0.4 + 1.98 + 16.66$$

$$= 19.38$$

Now: $Var(Q) = E(Q^2) - [E(Q)]^2 = 19.38 - (3.58)^2$

(by formula) $= 19.38 - 12.8164 = 6.5636$

General Info

Reminder: RVs

Discrete?

Terms & Notation

Summary

Discrete Random Variables

- Discrete RVs are RVs that will take one of a countable set of values.
- When working with a discrete random variable, it is common to need or use the RV's
    - probability distribution: the values the RV can take and their probabilities
    - probability function: a function where $f(x) = P(X = x)$
    - cumulative probability function: a function where $F(x) = P(X \leq x)$.
    - mean: a value for X defined by $EX = \sum_x x \cdot f(x)$
    - variance: a value for X defined by $VarX = \sum_x (x - \mu)^2 \cdot f(x)$

Your Turn:

Chapter 5 Handout 1

* Quiz II on Thursday.
  - Check sample quiz & note sheet on the course page.
  - Bring your calculator.
* HW 4 solution posted.
* HW 5 posted. Due Thursday March 5^{th}.

Common Distributions

Working with Off The Shelf Random Variables

General Info

Common Distributions

Background

Common Distributions

Why Are Some Distributions Worth Naming?

Even though you may create a random variable in a unique scenario, the way that it's probability distribution behaves (mathematically) may have a lot in common with other random variables in other scenarios. For instance,

> I roll a die until I see a 6 appear and then stop. I call X the number of times I have to roll the die in total.
>
> I flip a coin until I see heads appear and then stop. I call Y the number of times I have to flip the coin in total.
>
> I apply for home loans until I get accepted and then I stop. I call Z the number of times I have to apply for a loan in total.

Why Are Some Distributions Worth Naming? (cont)

In each ot the above cases, we count the number of times we have to do some action until we see some specific result. The only thing that really changes from the random variables perspective is the likelihood that we see the specific result each time we try.

Mathematically, that's not a lot of difference. And if we can really understand the probability behavior of one of these scenarios then we can move our understanding to the different scenario pretty easily.

By recognizing the commonality between these scenarios, we have been able to identify many random variables that behave very similarly. We describe the similarity in the way the random variables behave by saying that they have a common/shared distribution.

We study the most useful ones by themselves.

The Bernoulli Distribution

General Info

Common Distributions

Background

Bernoulli

The Bernoulli Distribution

Origin: A random experiment is performed that results in one of two possible outcomes: success or failure. The probability of a successful outcome is p.

Definition: X takes the value 1 if the outcome is a success. X takes the value 0 if the outcome is a failure.

probability function: probability of success

success

$$f(x) = \begin{cases} p & x = 1, \\ \underline{1-p} & x = 0, \\ 0 & o.w. \end{cases}$$

prob. of failure. Failure

which can also be written as

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & \boxed{x = 0, 1,} \\ 0 & o.w. \end{cases}$$

i.e. for success ; $x=1 \rightarrow f(\text{success}) \equiv f(1) = P(X=1) = p^1(1-p)^{1-1} = p$

for failure ; $x=0 \rightarrow f(\text{failure}) = f(0) = P(X=0) = p^0(1-p)^{1-p} = 1-p$

Bernoulli Distribution

Expected Value and Variance

General Info

Common Distributions

Background

Bernoulli

The Bernoulli Distribution

Expected value: $E(X) = p$

proof: (optional reading):

$$E(X) = \sum_{x=0}^{1} x\,p(x) = 0\,p(0) + (1)\,p(1)$$

$$= 0 + (1)\,p$$

$$= p$$

Recall: in Bernoulli: $p(0) = 1-p$

$p(1) = p$

General Info

Common Distributions

Background

Bernoulli

The Bernoulli Distribution

Variance: $Var(X) = (1-p) \cdot p$

Proof: (optional reading):

First find $E x^2$. Then use the formula

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

$$E(X^2) = \sum_{x=0}^{1} x^2 f(x) = 0^2 f(0) + 1^2 f(1)$$

$$= 0 + f(1)$$

$$= p$$

$$\Rightarrow Var(X) = E(X^2) - [E(X)]^2$$

$$= p - p^2 = p(1-p) \quad //$$

General Info

Common Distributions

Background

Bernoulli

The Bernoulli Distribution

A few useful notes:

- In order to say that " X has a bernoulli distribution with success probability p " we write $X \sim Bernoulli(p)$
- Trials which results in which the only possible outcomes are "success" or "failure" are called **Bernoulli Trials**
- The value p is the Bernoulli distribution's **parameter**. We don't treat parameters like random values - they are fixed, related to the real process we are studying.
- "Success" does not mean something we would perceive as "good" has happened. It just means the outcome we were watching for was the outcome we got.
- Please note: we have two outcomes, but the probability for each outcome is **not** the same (duh!).

outcomes
success → with prob. p
failure → with prob. $(1-p)$

The Binomial Distribution

The Binomial Distribution → Repeat Bernoulli trial "n" times.

Origin: A series of n independent random experiments, or trials, are performed. Each trial results in one of two possible outcomes: success[redacted] or failure. The probability of a successful outcome, p, is the same across all trials.

Definition: For n trials, X is the number of trials with a successful outcome. X can take values $0, 1, \ldots, n$.

probability function:

With $0 < p < 1$,

$$P(X=x) = f(x) = \begin{cases} \dfrac{n!}{x!(n-x)!} p^x (1-p)^{n-x} & x = 0, 1, \ldots, n \\ 0 & \underbrace{o.w.}_{\text{other wise}} \end{cases}$$

`

where $n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1$ and $0! = 1$.

If we have only $n=1$ trial $\Rightarrow f(x) = \underbrace{\dfrac{1!}{x!\,(1-x)!}}_{=1} p^x (1-p)^{1-x} = p^x (1-p)^{1-x} \sim \text{Bernoulli}(p)$