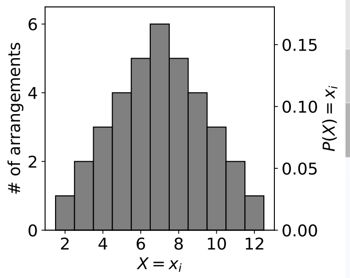
EGM101 – Skills Toolbox

Week 7, Part 4: Discrete Probability Distributions

Roll two D6

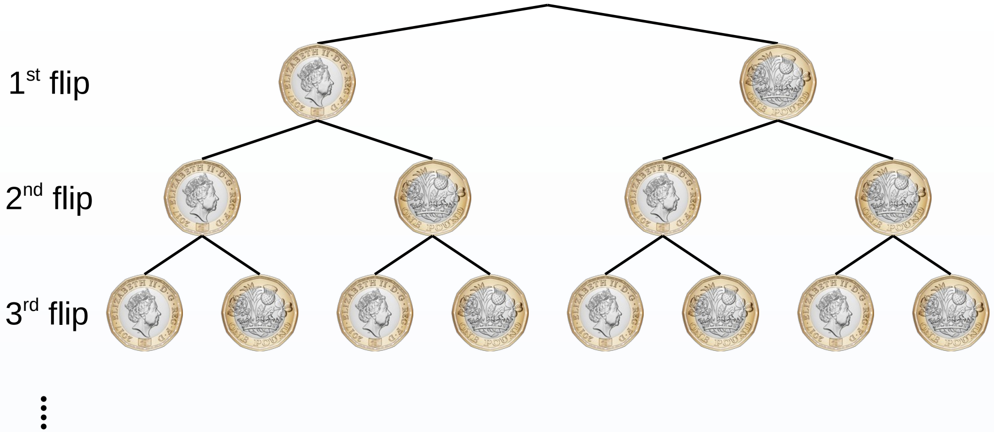
- If we roll two fair (D6) dice, what is the probability of getting a particular number?
  - i.e., outcome X is equal to some discrete value
- Probability distribution: a description of the probability of different events
- We can use probability distributions to estimate the likelihood of:
  - A particular outcome (e.g., rolling a 6)
  - A range of outcomes (e.g., rolling > 9)

$X = x$	Possible arrangements	# arrangements	$P(X = x)$
2	⚀⚀	1	1 / 36
3	⚀⚁ ⚁⚀	2	2 / 36
4	⚀⚂ ⚂⚀ ⚁⚁	3	3 / 36
5	⚀⚃ ⚃⚀ ⚁⚂ ⚂⚁	4	4 / 36
6	⚀⚄ ⚄⚀ ⚁⚃ ⚃⚁ ⚂⚂	5	5 / 36
7	⚀⚅ ⚅⚀ ⚁⚄ ⚄⚁ ⚂⚃ ⚃⚂	6	6 / 36
8	⚁⚅ ⚅⚁ ⚂⚄ ⚄⚂ ⚃⚃	5	5 / 36
9	⚂⚅ ⚅⚂ ⚃⚄ ⚄⚃	4	4 / 36
10	⚃⚅ ⚅⚃ ⚄⚄	3	3 / 36
11	⚄⚅ ⚅⚄	2	2 / 36
12	⚅⚅	1	1 / 36
Total		36	1



Successive coin flips

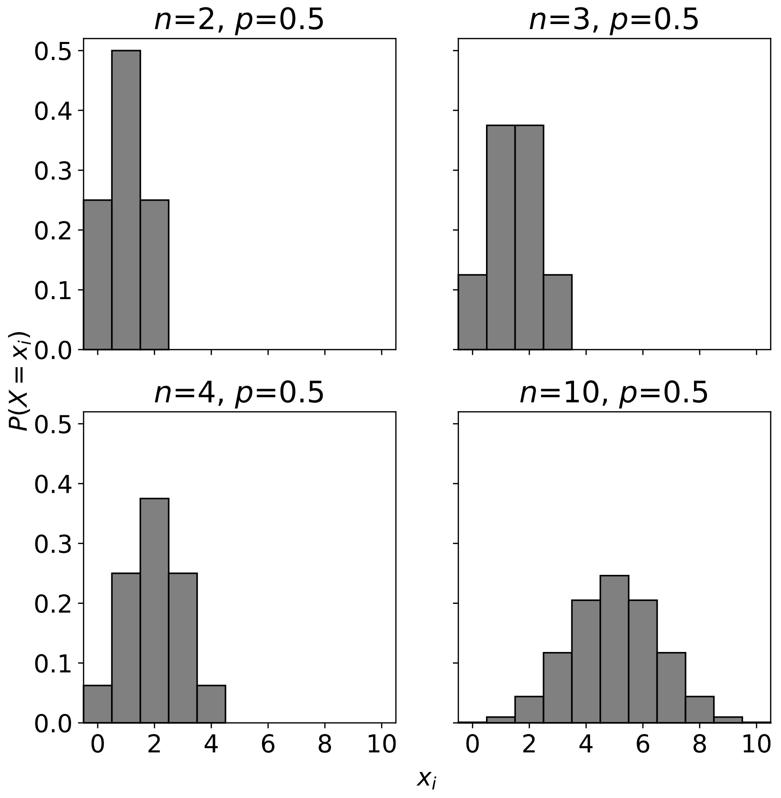
- Take a fair coin, flip it n times
  - (alternatively, flip n coins)
- Question: after n flips, how many ways can we end up with:
  - n heads?
  - $n - 1$ heads?
  - …
  - 0 heads?
- This is an example of a binomial experiment:
  - Fixed number of n trials
  - Two possible outcomes (“success”, “failure”): Bernoulli trial
  - Independent, identical trials (i.e., probability remains the same each time)



Number of heads, m							$P(m)$				
Flips	0	1	2	3	4	Total	0	1	2	3	4
1	1	1				2	1/2	1/2			
2	1	2	1			4	1/4	2/4	1/4		
3	1	3	3	1		8	1/8	3/8	3/8	1/8	
4	1	4	6	4	1	16	1/16	4/16	6/16	4/16	1/16

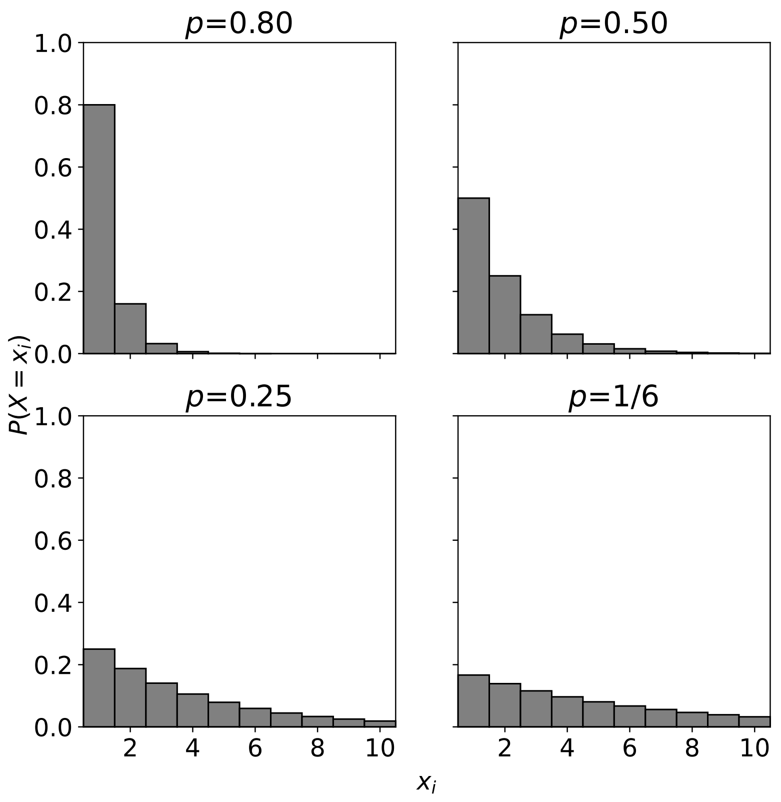
The Binomial Distribution

- Outcome of a binomial experiment, X, follows a binomial probability distribution
- Probability that X is some $x_i \leq n$ depends on:
  - Probability of success, p
  - Number of trials, n
- Note that as n increases, "peak" probability decreases:
  - Remember: sum of individual probabilities must equal 1
  - In effect, total probability is spread over a larger range of possible x values



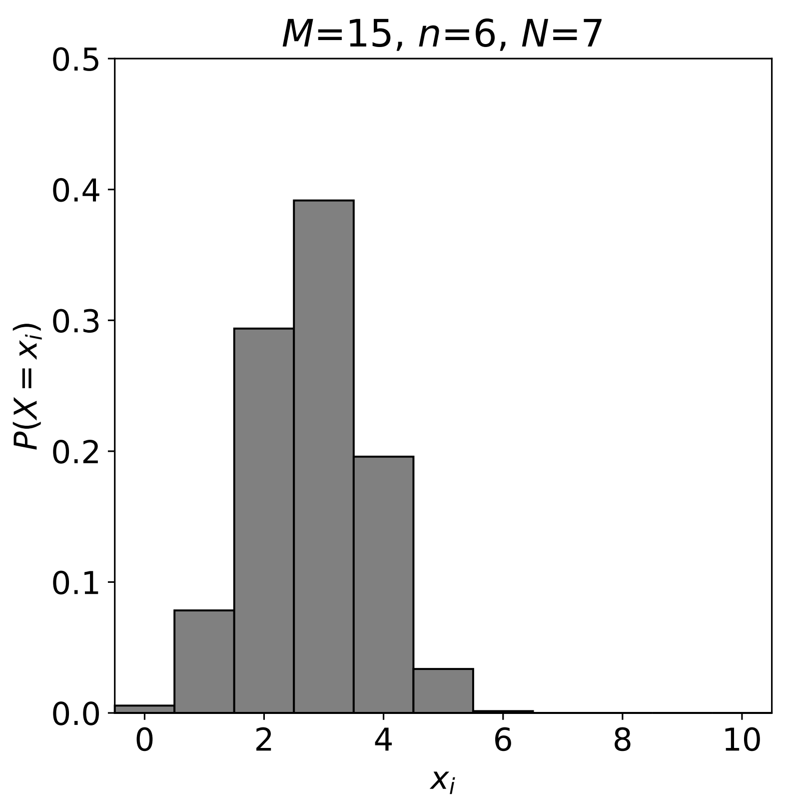
The Geometric Distribution

- What is the probability that we don't get a head until we've flipped n coins?
- Similar setup to the binomial distribution:
  - Bernoulli trial ("success", "failure")
  - Independent trials
- Probability that the first success is on the k^{th} trial depends on:
  - Probability of success, p
- Note that probability drops quickly for p closer to 1



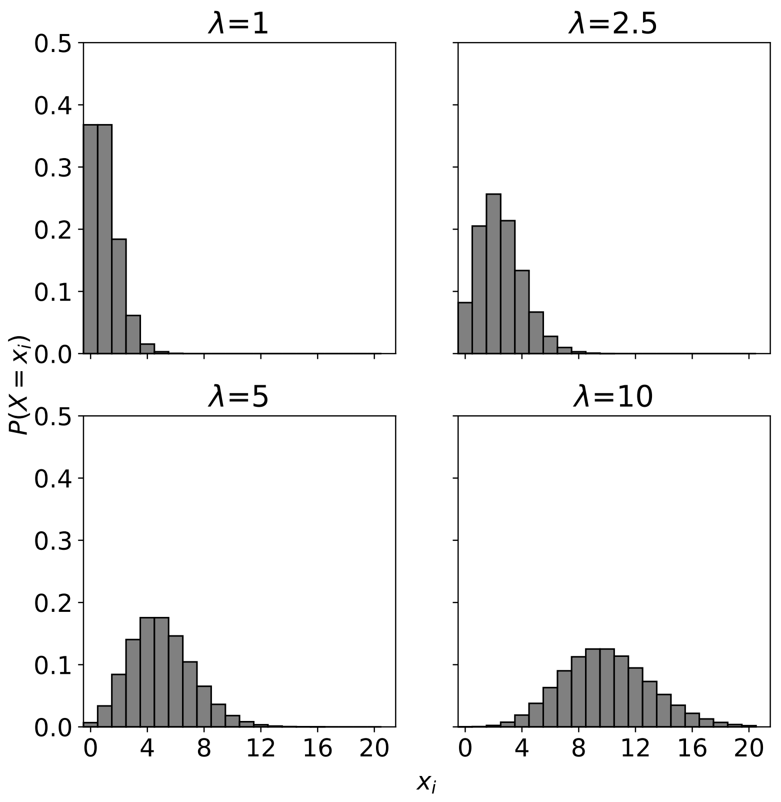
The Hypergeometric Distribution

- If we have a mixture of marbles (9 red, 6 blue), and we select 7 marbles, what is the probability that we end up with x blue marbles?
  - Hypergeometric experiment
- Sampling without replacement:
  - Not independent
  - Not Bernoulli trials
- Probability depends on:
  - Size of the whole group, M
  - Size of the "group of interest", n
  - Number of samples taken, N



The Poisson Distribution

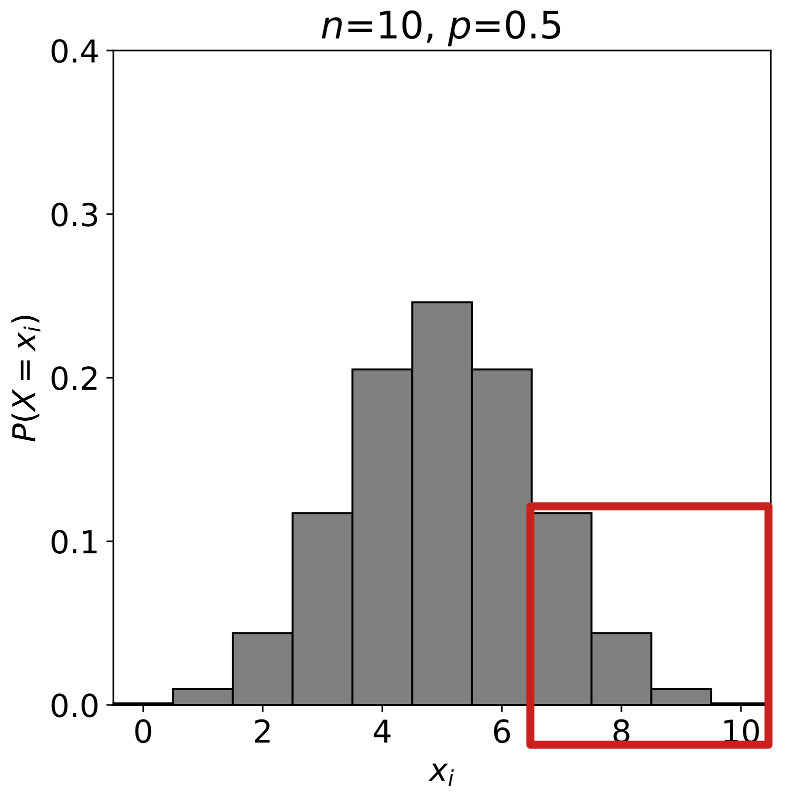
- If we catch 10 salmon per hour, how many can we expect to catch in the next 15 minutes?
- Poisson distribution: probability of x events occurring in a fixed interval
  - Assumes that events are independent
  - Depends on average rate of occurrence, λ
- Peak probability occurs close to λ
  - As λ increases, so does spread
  - Technically unbounded



More on probability distributions

- Can use distributions to calculate probability of range of events
- Example: Flip a fair coin 10 times.
- What is the probability of getting more than 7 heads?
  - i.e., what is $P(X \geq 7)$?

$$P(X \geq 7) = P(X=7) + P(X=8) + P(X=9) + P(X=10)$$



Summary

- Probability distributions help us understand the likelihood of a given outcome/event
- Theoretical distributions save us from having to calculate probability of each individual outcome/event
- With a distribution, we can also calculate probability of a range of events/outcomes

Additional Resources

- Illowsky and Dean, Chapter 4
- Caswell, Chapter 13
- Weiss, Chapter 5.4
- The mathematical secrets of Pascal's triangle [TED-Ed]
- Pascal's Triangle [Numberphile]
- 14 Super Bowl Coin Tosses [Numberphile]