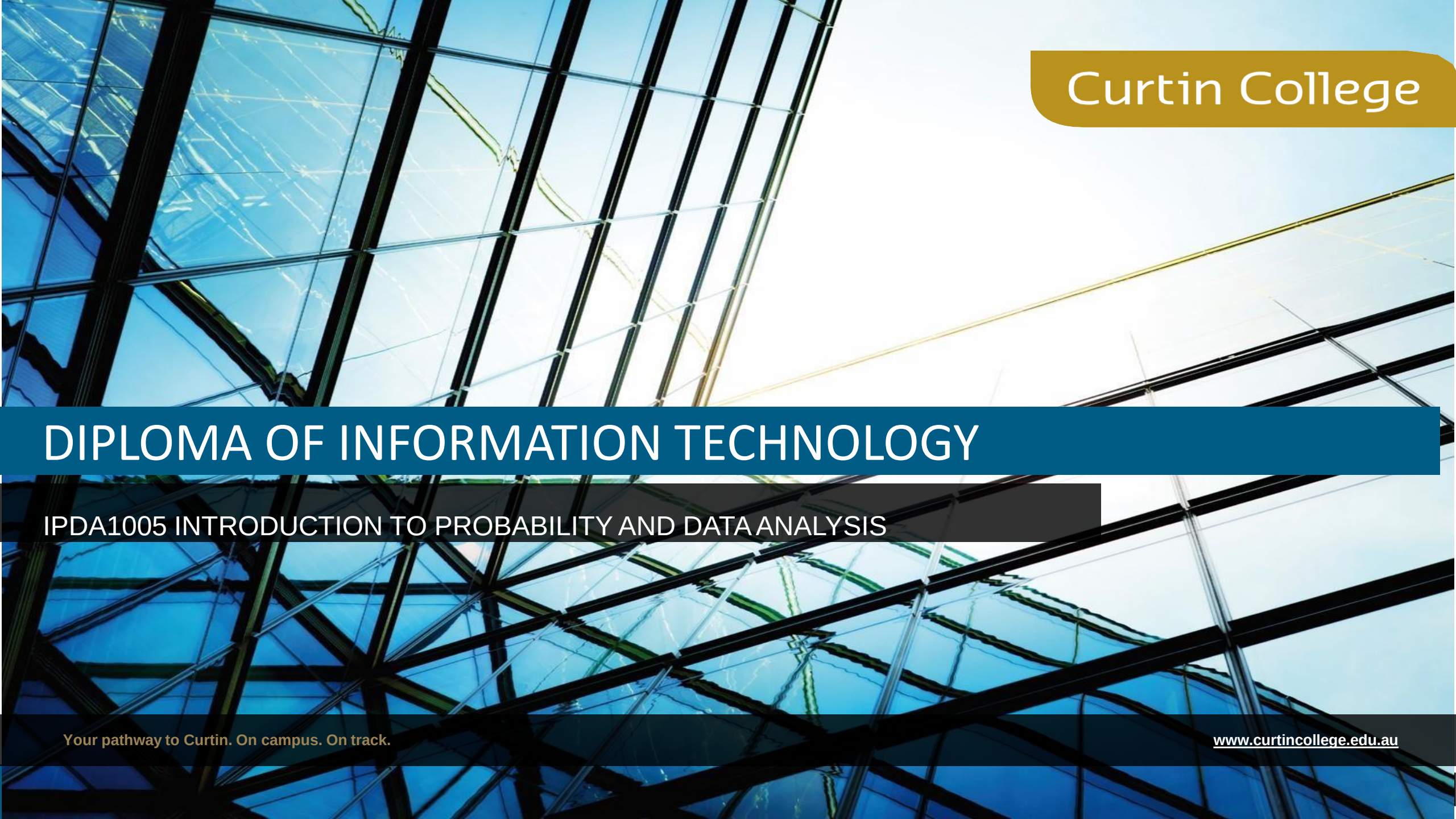


DIPLOMA OF INFORMATION TECHNOLOGY

IPDA1005 INTRODUCTION TO PROBABILITY AND DATA ANALYSIS

Acknowledgement

We respectfully acknowledge the Elders and custodians of the Whadjuk Nyungar nation, past and present, their descendants and kin. Curtin College Bentley Campus enjoys the privilege of being located in Whadjuk / Nyungar Boodjar (country) on the site where the Derbal Yerrigan (Swan River) and the Djarlgarra (Canning River) meet. The area is of great cultural significance and sustains the life and well being of the traditional custodians past and present.

Outline

Discrete Uniform Distribution

- If a random variable X takes n possible values, $a_1, a_2, \ldots, a_n$, all with probability $\frac{1}{n}$, then X is said to have *discrete uniform* distribution on $\{a_1, \ldots, a_n\}$.
- **Example:** Roll a die and let X be the number that came up. If the die is well balanced, then X has discrete uniform on $\{1, 2, 3, 4, 5, 6\}$.
- **Example:** A box has 9 balls numbered

$$2, \quad 3, \quad 5, \quad 7, \quad 11, \quad 13, \quad 17, \quad 19, \quad 23,$$

i.e., the prime numbers below 25. A ball is chosen at random and let X be the number that came up.

Then X has discrete uniform on $\{2, 3, 5, 7, 11, 13, 17, 19, 23\}$.
$P(X = x) = \frac{1}{9}$ for $x \in \{2, 3, 5, 7, 11, 13, 17, 19, 23\}$.

Expectation and Variance of Discrete Uniform

Result:

1. Let $X \sim$ Discrete Uniform $\{1, 2, \ldots, n\}$ so that $p(x) = \frac{1}{n}$ for $x = 1, 2, \ldots n$. Then

$$E(X) = \frac{n+1}{2} \quad \text{and} \quad Var(X) = \frac{n^2 - 1}{12}$$

2. Let $X \sim$ Discrete Uniform $\{0, 1, 2, \ldots, n\}$ so that $p(x) = \frac{1}{n+1}$ for $x = 0, 1, 2, \ldots n$. Then

$$E(X) = \frac{n}{2} \quad \text{and} \quad Var(X) = \frac{n(n+2)}{12}$$

Expectation and Variance of Discrete Uniform

- **Proof:** For discrete uniform on $\{1, 2, \ldots, n\}$:

$$\begin{aligned} E(X) &= \sum_x xp(x) \\ &= \frac{1}{n}\sum_{x=1}^{n} x \\ &= \frac{1}{n}\left(\frac{n\,(n+1)}{2}\right) \\ &= \frac{n+1}{2} \end{aligned}$$

Expectation and Variance of Discrete Uniform

$$
\begin{aligned}
E(X^2) &= \sum_x x^2 p(x) \\
&= \frac{1}{n}\sum_{x=1}^{n} x^2 \\
&= \frac{1}{n}\left(\frac{n\,(n+1)\,(2n+1)}{6}\right) \\
&= \frac{(n+1)\,(2n+1)}{6}.
\end{aligned}
$$

$$
\begin{aligned}
\text{Hence:} \quad Var(X) &= \frac{(n+1)\,(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 \\
&= \frac{n^2-1}{12}.
\end{aligned}
$$

Expectation and Variance of Discrete Uniform

- The proof for discrete uniform on $\{0, 2, \ldots, n-1\}$ is left as an exercise.
- **Example:** Let X be the number that came up when a die is rolled once. Here X has uniform distribution over the set $\{1, 2, 3, 4, 5, 6\}$.
- Hence $E(X) = \frac{6+1}{2} = 3.5$ and $Var(X) = \frac{6^2-1}{12} = \frac{35}{12}$.
- **Exercise:** Find the mean and standard deviation for the prime number example. (Note: You cannot use the formulas for expectation and variance of discrete uniform in this case.)

Bernoulli random variable

A random variable X that takes just two values 0 and 1 such that $P(X = 1) = p$ is a *Bernoulli* random variable. Its distribution is referred to as a Bernoulli(p) distribution, denoted by $X \sim Ber(p)$.

- Bernoulli is a *family* of distributions, with different distributions for different values of p. If $X \sim Ber(p)$, then $P(X = 1) = p$ and $P(X = 0) = 1 - p$.
- We can write the PMF of X as

$$P(X = x) = p^x(1 - p)^{1-x}, \qquad x = 0, 1$$

- Examples:
  - Result of flipping a coin
  - Does a driver fail a breathalyzer test?
  - Does an individual have a particular disease?
- Any experiment with two possible outcomes is a *Bernoulli trial*, and the result is a Bernoulli RV.

Bernoulli Trials - examples

- Let X be the number of heads obtained when a coin whose probability of heads is α is tossed. Then $X \sim Ber(\alpha)$.
- Let Y be the number of tails obtained when a fair coin is tossed. Then $Y \sim Ber(0.5)$.
- Let W be the number of white balls obtained when a ball is chosen at random from an urn containing 7 white balls and 3 black balls. Then $W \sim Ber(0.7)$.
- Let T take value 1 or 0 depending on whether a six is obtained when a fair die is rolled. Then $T \sim Ber(\frac{1}{6})$.
- Let Z be the number of people with Covid-19 when we randomly select a person from a population where the prevalence of Covid-19 is 1.2%. Then $Z \sim Ber(.012)$.

Bernoulli Distribution - Mean and Variance

- If $X \sim Ber(p)$, $E(X) = 0(1-p) + 1(p) = p$.
- Similarly, $E(X^2) = p$
- Thus $Var(X) = p - p^2 = p(1-p)$.
- **Example:** Let X the number of heads obtained when a fair coin is tossed once. Then

$$E(X) = \frac{1}{2} \quad \text{and} \quad Var(X) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}$$

- **Example:** Let X take value 1 if the roll of a fair die results in a six, and 0 otherwise. Then

$$E(X) = \frac{1}{6} \quad \text{and} \quad Var(X) = \left(\frac{1}{6}\right)\left(\frac{5}{6}\right) = \frac{5}{36}$$

Binomial Distribution

If X is the number of successes in a sequence of n independent Bernoulli trials with probability of success p, then X has a *binomial* distribution with parameters n and p. We write this as

$$X \sim Bin(n, p)$$

- **Examples:**
  - A coin is tossed 20 times and the number of heads is counted.
  - A NAS server contains 10 hard drives and the probability of any one drive failing in the first year is p and is independent of the failure of any other drives. The quantity of interest is the number of failed drives in the first year.
  - Thirty monkeys infected with HIV are administered a new drug. We want to know how many monkeys show a marked improvement in immunological response.

Binomial distribution - probabilities

Consider $X \sim Bin(n, p)$:

- For X to take the value k, the string of results from the n Bernoulli trials must have exactly k Ss (successes) and $n - k$ Fs (failures).
- The chance of this happening for *any particular string* of Bernoulli trials is $p^k(1-p)^{n-k}$.
- Since there are k positions out of n in which Ss can occur, there are $\binom{n}{k}$ such strings.
- Thus $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.
- In general, the binomial probability formula is given as follows

If $X \sim Bin(n, p)$, then

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$$

for $x = 0, 1, \ldots, n$.

Binomial distribution - example

In the HIV-infected monkeys example, if the chance of improvement for an individual monkey is 60%, find the probability that

- exactly 15 monkeys showed improvement
- at most 15 monkeys showed improvement
- more than 15 monkeys showed improvement

Let X be the number of monkeys showing improvement. $X \sim Bin(30, .6)$

- $P(x = 15) = \binom{30}{15}.6^{15}.4^{15} = 0.0783$
- $P(X \leq 15) = \sum_{k=0}^{15} \binom{30}{k}.6^{k}.4^{30-k} = 0.1754$
- $P(X > 15) = 1 - P(X \leq 15) = 1 - 0.1754 = 0.8246$

  The last two calculations could be quite tedious as they involve summing 16 and 15 probabilities. Note how the CDF reduces the work. Printed tables may be available for n up to 30.

```
> dbinom(15,30,.6)
[1] 0.07831221
```

```
> pbinom(15,30,.6)
[1] 0.1753691
```

Binomial distribution - Mean and Variance

If $X \sim Bin(n, p)$, then

$$E(X) = np \quad \text{and} \quad Var(X) = np(1-p)$$

- **Example:** Let X the number of heads obtained when a fair coin is tossed 100 times. Find the mean and variance of X.

$$E(X) = 100\left(\frac{1}{2}\right) = 50 \quad \text{and} \quad Var(X) = 100\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = 25$$

- **Example:** Let X be the number of 2's in 45 rolls of a fair die. Find the mean and standard deviation of X.

$$E(X) = 45\left(\frac{1}{6}\right) = 7.5 \quad \text{and} \quad Var(X) = 45\left(\frac{1}{6}\right)\left(\frac{5}{6}\right) = 6.25$$

so $SD(X) = 2.5$.

Geometric Distribution

- Suppose Bernoulli trials with probability of success p are conducted until a success is obtained. If X is the number of failures before the success, then X follows the *geometric* distribution.
- E.g., for the outcome $FFFFS$, $X = 4$ and the probability is $p(1-p)^4$.

The number of failures X before the first success in a sequence of Bernoulli trials with success probability p follows a geometric(p) distribution, often written as $X \sim Geom(p)$. The PMF is given by

$$P(X = x) = p(1-p)^x$$

for $x = 0, 1, 2, \ldots$

Geometric Distribution - properties

- Geometric distribution is an infinite discrete distribution.
- Its probabilities follow a geometric series, hence the name.
- Therefore the geometric distribution CDF has the simple formula: $P(X \leq k) = 1 - q^{k+1}$, where $q = 1 - p$.
- **Proof:** Recall that the sum of a geometric series with first term a and common ratio r is given by $S_n = \frac{a(1-r^n)}{1-r}$.

$$
\begin{aligned}
P(X \leq k) &= \sum_{x=0}^{k} p(1-p)^x \\
&= \sum_{x=0}^{k} pq^x \quad \text{using } q \text{ for } 1-p. \\
&= \frac{p(1-q^{k+1})}{1-q} \\
&= 1 - q^{k+1}
\end{aligned}
$$

Geometric Distribution - example

Example: Let X have $Geom(.6)$ distribution. Find

1. $P(X = 4)$
2. $P(X > 2)$
3. $P(X \leq 4)$

Solution:

1. $P(X = 4) = pq^4 = .6 \times .4^4 = 0.01536$
2. $P(X > 2) = 1 - P(X \leq 2) = 1 - (1 - q^3) = .4^3 = 0.064$
3. $P(X \leq 4) = 1 - q^5 = 1 - .4^5 = 0.98976$

Examples

Example: A fair die is rolled until a six is obtained. Find the probability that the number of rolls needed is

1. exactly 4
2. at least 3
3. at most 3

Solution: Let $X \sim Geom\left(\frac{1}{6}\right)$

1. $P(\text{exactly 4 rolls}) = P(X = 3) = \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^3 = .09645$
2. $P(\text{at least 3 rolls}) = P(X \geq 2) = 1 - P(X \leq 1) = \left(\frac{5}{6}\right)^2 = .69444$
3. $P(\text{at most 3 rolls}) = P(X \leq 2) = 1 - \left(\frac{5}{6}\right)^3 = .4213$

Geometric Distribution - Mean and Variance

If $X \sim Geom(p)$,

$$E(X) = \frac{1-p}{p} = \frac{q}{p} \quad \text{and} \quad Var(X) = \frac{1-p}{p^2} = \frac{q}{p^2}$$

- **Example:** A Bernoulli trial with success probability 0.2 is repeated until a success is obtained. Let X be the number of failures. Find the mean and variance of X.

  Solution: $X \sim Geom(.2)$, so
  $E(X) = \frac{.8}{.2} = 4$ and $Var(X) = \frac{.8}{.2^2} = 20$.

- **Example:** A fair die is rolled until a six is obtained. Find the mean and variance of the number of rolls.

  Solution: Let $X =$ no. of rolls, and $Y = X - 1$, so $Y \sim Geom(\frac{1}{6})$.
  $E(Y) = \frac{5/6}{1/6} = 5$ and $Var(Y) = \frac{5/6}{1/36} = 30$.
  As $X = Y + 1$, $E(X) = E(Y) + 1 = 6$ and $Var(X) = Var(Y) = 30$

Negative Binomial Distribution

We can generalise the geometric distribution by considering stopping after r successes where r is a positive integer. This gives rise to the *negative binomial* distribution.

The number of failures X before obtaining r successes in a sequence of independent Bernoulli trials with success probability p follows a negative binomial(r, p) distribution, often written as $X \sim NB(r, p)$.
The PMF is given by

$$P(X = x) = \binom{x + r - 1}{r - 1} p^r (1 - p)^x$$

for $x = 0, 1, 2, \ldots$

- Note that the r successes don't have to be consecutive. If $r = 3$, $FFSFSS$ is an acceptable outcome.
- When $r = 1$, this becomes the geometric distribution. Note that the PMF of $NB(1, p)$ is the same as that of $Geom(p)$.

Negative Binomial Distribution - Examples

Example: A Bernoulli trial with a success probability of 0.4 is repeated until 2 successes are obtained. Find the probability that the number of failures is

1. exactly 3
2. at most 3
3. at least 3

Solution: Let X be the number of failures. Then $X \sim NB(2, .4)$.

1. $P(X = 3) = \binom{3+2-1}{2-1}.4^2.6^3 = .13824$
2. $P(X \leq 3) = .66304$
3. $P(X \geq 3) = 1 - P(X \leq 2) = 1 - .5248 = .4752$

Negative Binomial Distribution - Examples

- **Example:** A fair coin is tossed until 3 heads are obtained. Find the probability that the number of tosses needed is
  1. exactly 5
  2. at most 5
  3. at least 5

  Solution: Let X be the number of tails in the sequence. Then X is a $NB\,(3, .5)$ random variable.
  1. $P(\text{exactly 5 tosses}) = P(X = 2) = \binom{2+3-1}{3-1}.5^3.5^2 = 6 \times .5^5 = .1875$
  2. $P(\text{at most 5 tosses}) = P(X \leq 2) = .5$
  3. $P(\text{at least 5 tosses}) = P(X \geq 2) = 1 - P(X \leq -1) = .6875$

Negative Binomial Distribution - Mean and Variance

If $X \sim NB(r, p)$,

$$E(X) = \frac{r(1-p)}{p} = \frac{rq}{p} \quad \text{and} \quad Var(X) = \frac{r(1-p)}{p^2} = \frac{rq}{p^2}$$

Example: A Bernoulli trial has a success probability of 0.2. The trial is repeated until 3 successes are obtained. Find the mean and variance of the number of failures.

Solution: $X \sim NB(3, 0.2)$,

so $E(X) = \frac{3(.8)}{.2} = 12$ and $Var(X) = \frac{3(.8)}{.2^2} = 60$.

Example: A fair die is rolled until four 6's are obtained. Find the mean and variance of the number of rolls.

Solution: Let Y = no. of failures, and $Y = X - 4$. Then

$Y \sim NB(4, \frac{1}{6})$, $E(Y) = \frac{4(5/6)}{1/6} = 20$ and $Var(Y) = \frac{4(5/6)}{1/36} = 120$.

$X = Y + 4$, so $E(X) = E(Y) + 4 = 24$ and $Var(X) = Var(Y) = 120$

Hypergeometric Distribution

Regarding the probability model for the number of successes obtained when sampling from a finite dichotomous population:

- the binomial distribution is the appropriate model when sampling *with replacement*
- the *hypergeometric* distribution is the appropriate model when sampling *without replacement*.

For example, for the probability of turning up two aces from a standard deck when ten cards are selected, -

- use the binomial distribution if selection is with replacement.
- use the hypergeometric distribution if selection is without replacement.

Hypergeometric Distribution

- We use the following general scenario and notation.
  - The population to be sampled consists of N items.
  - Each item can be characterized as a success S or a failure F, and there are M successes in the population.
  - A sample of n items is selected without replacement in such a way that each subset of size n is equally likely to be chosen.
- It does not matter whether we sample the n items one by one without replacing, or sample all n items in one go.
- The random variable of interest is the number X of successes in the sample. The probability distribution of X depends on the parameters N, M and n.

Hypergeometric Distribution

The number of successes X when n items are sampled without replacement from a population of N items where M items are marked as success follows a hypergeometric(n, M, N) distribution, often written as $X \sim HG(n, M, N)$. The PMF is given by

$$P(X = x) = h(x; n, M, N) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

for integer x satisfying $\max(0, n - N + M) \leq x \leq \min(n, M)$.

Hypergeometric Distribution - example

- An urn contains 6 white balls and 5 black balls. We sample $n = 5$ balls without replacement. What is the probability that the number of white balls in the sample is 3?
- If we define X to be the number of white balls in the sample, $X \sim HG(5, 6, 11)$.
- $P(X = x) = \frac{\binom{6}{x}\binom{5}{5-x}}{\binom{11}{5}}$ for integer x in $[0, 5]$.
- $P(X = 3) = \frac{\binom{6}{3}\binom{5}{2}}{\binom{11}{5}} = \frac{200}{462} \approx 0.4329.$
- Here is the PMF for all values of X.

x	0	1	2	3	4	5
$p(x)$	$\frac{1}{462}$	$\frac{30}{462}$	$\frac{150}{462}$	$\frac{200}{462}$	$\frac{75}{462}$	$\frac{6}{462}$

Hypergeometric Distribution - Mean and Variance

For $X \sim HG(n, M, N)$, the mean and variance are given by

$$E(X) = n\left(\frac{M}{N}\right), \quad Var(X) = n\left(\frac{M}{N}\right)\left(1 - \frac{M}{N}\right)\left(\frac{N-n}{N-1}\right)$$

- If we denote the population proportion of successes $\frac{M}{N}$ by p, then the formulae become -
  $E(X) = np, \quad Var(X) = np\,(1-p)\left(\frac{N-n}{N-1}\right)$
- Recall that for $Bin(n, p)$ the respective formulae are
  $E(X) = np, \quad Var(X) = np\,(1-p)$
- The extra factor $\frac{N-n}{N-1}$ in the hypergeometric variance is called the *finite population correction (FPC)*. FPC is always < 1. For fixed sample size n, FPC approaches 1 as $N \to \infty$.
- These results indicate that the hypergeometric distribution is similar to the binomial distribution but has a smaller variance.

Hypergeometric distribution - Capture-recapture Models

Suppose a national park has a population of N animals of a particular species. At a given time, M are captured, tagged and released. Later, a random sample of n animals are (re)captured.

- Let X be the number of tagged animals in the second sample.
- The distribution of X will be hypergeometric(n, M, N). Initially, N is often unknown.

Exercise: Five individuals from an endangered local animal population have been trapped, tagged, and released. After they have had an opportunity to mix, ten of these animals are trapped at random. Let X be the number of tagged animals in the second sample. If there are actually 25 animals of this type in the region, what is the probability that -

1. $X = 2$?
2. $X \leq 2$?

Hypergeometric distribution - Capture-recapture Models

Thirty animals from a population thought to be near extinction in a region have been caught, tagged, and released to mix into the population. After they have had an opportunity to mix, a random sample of 20 of these animals is trapped and 5 of them are found to be tagged. Use this to estimate the population of these animals in the region.

Solution: The proportion of tagged animals in the population is estimated by the sample proportion $\frac{5}{20}$. This gives us $\frac{30}{N} \approx \frac{5}{20}$, from which we get the estimate $N \approx \frac{30 \times 20}{5} = 120$.

Poisson Distribution

- The Poisson distribution commonly arises where we count the number of events occurring in interval of time or region of space (but *not* in a known number of cases).
- Examples are the number of severe tropical cyclones making landfall per season, the number of emails you receive per hour, the number of accidents occurring in one kilometre stretch of a highway, the number of customer enquiries received in a day, etc.
- The Poisson distribution has a single parameter λ, interpreted as the rate of occurrence of the events that are being modelled.

If $X \sim Pois(\lambda)$, then its PMF is

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, 3 \ldots$$

- Because $e^{\lambda} = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$, it follows that $\sum_{x=0}^{\infty} p(x) = 1$.

Poisson Distribution - Mean and Variance

If $X \sim Pois(\lambda)$, $E(X) = Var(X) = \lambda$.

Example: A telephone switchboard averages 6 calls per hour. Assuming a Poisson distribution, find -

1. the probability that exactly five calls come through in a one hour period
2. the chance that at least five calls come through between 10:00 and 11:00
3. expectation and variance of the number of calls that come through between 11:00 and 11:15
4. the probability that no more than one call comes through between 11:00 and 11:15

Poission distribution - example solution

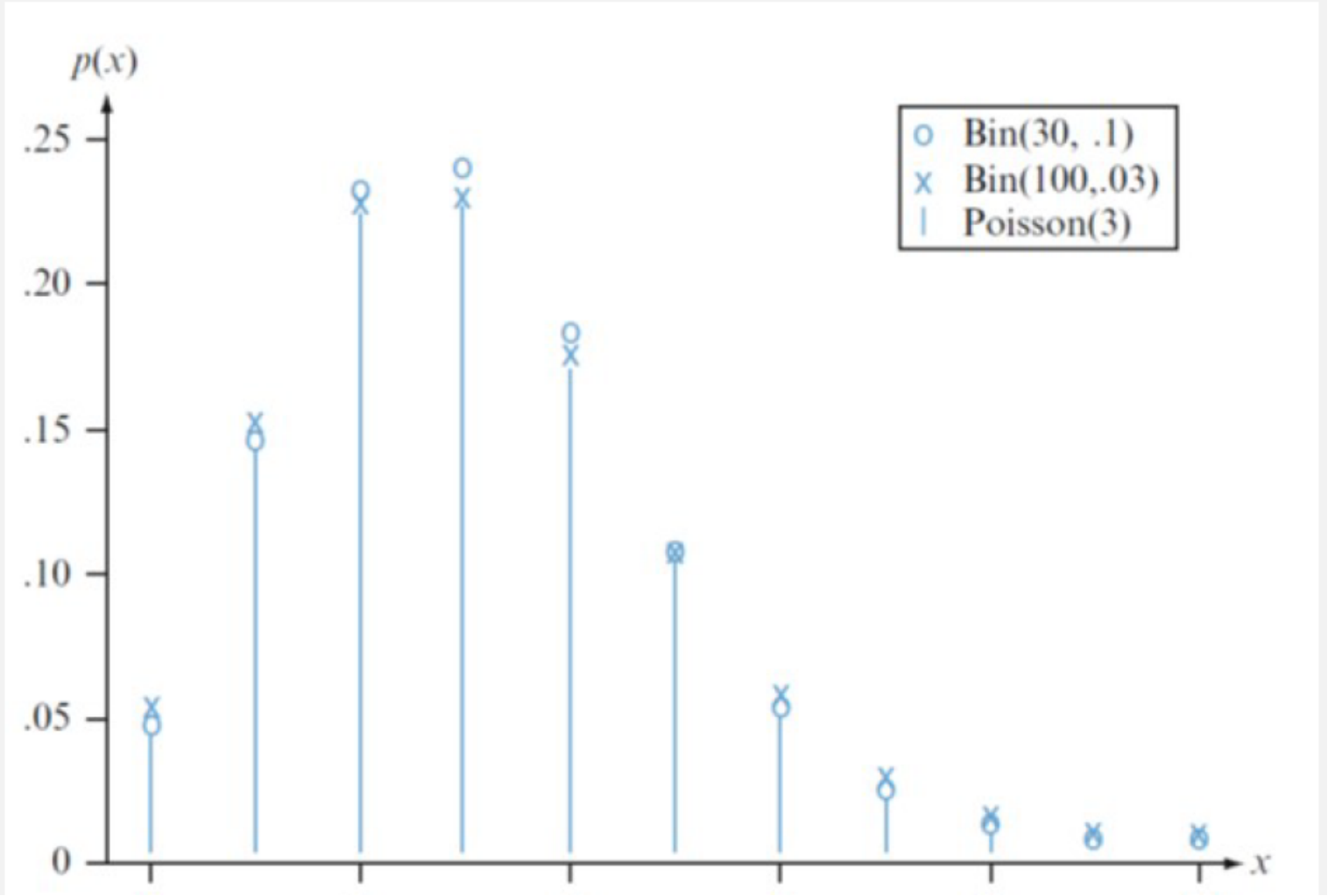
1. If X is the number of phone calls that come through in one hour, $X \sim Pois(6)$ and $P(X = 5) = \frac{e^{-6}6^5}{5!} = 0.1606$
2. $P(X \geq 5) = 1 - P(X \leq 4) = 1 - .285057 = .7149$
3. In a quarter hour period, the arrival rate is 1.5, so if Y is the number of calls that come through between 11:00 and 11:15, then $Y \sim Pois(1.5)$. So $E(Y) = Var(Y) = 1.5$
4. $P(Y < 1) = .557825$

Poisson Approximation of Binomial

Suppose that in the binomial pmf $b(x; n, p)$ we let $n \to \infty$ and $p \to 0$ in such a way that np approaches $\lambda > 0$. Let $p(x; \lambda)$ denote $P(X = x)$ where $X \sim Poi(\lambda)$. Then $b(x; n, p) \to p(x; \lambda)$. In practice, the approximation can be safely used if $n \geq 30$ and $np \leq 5$.

- In other words, if $X_n \sim Bin(n, p)$ and $np \to \lambda > 0$, then $P(X_n = x) \to \frac{e^{-\lambda}\lambda^x}{x!}$.
- We will compare the three distributions $Bin(30, 0.1)$, $Bin(100, 0.03)$ and $Pois(3)$. All of them have the same mean $\mu = 3$.
- According to the statement above, both $Bin(30, 0.1)$ and $Bin(100, 0.03)$ can be approximated by $Pois(3)$, but the approximation for $Bin(100, 0.03)$ is likely to be better.

Poisson Approximation of Binomial



Poisson Approximation of Binomial

- **Example:** The owner of a website is studying the distribution of the number of visitors to the site. Each day a million people independently decide whether to visit the site, with probability $p = 2 \times 10^{-6}$ of visiting. Use the Poisson approximation to provide an estimate of the probability of getting at least three visitors on a particular day.

  Solution If X is the number of visitors to the site,

  $$X \sim Bin(10^6, 2 \times 10^{-6}) \approx Poi(2)$$

  $$P(X \geq 3) = 1 - P(X \leq 2) = 1 - .6767 = .3233$$

Poisson Process

Consider a collection of random variables $\{N(t) : t \in [0, \infty)\}$, where $N(t)$ is the number of events occurring in the time interval $[0, t]$.
The $N(t)$ are said to be a *Poisson process* with rate λ if

- $N(0) = 0$
- $N(t) - N(s) \sim Pois\left(\lambda(t - s)\right)$ for $s < t$
- $N(t) - N(s)$ is independent of $N(s)$

Fact: If $X_1, X_2, X_3 \ldots$ are the inter-event times (X_1 is the time elapsed before the first event, X_2 is the time between the first and second events, etc.), then $X_1, X_2, X_3 \ldots$ are independent and have $Exp(\lambda)$ distribution.

Poisson process - example

The number of customers arriving at a 24-hour service facility follows a Poisson process with rate $\lambda = 3.5$ per hour.

1. Find the expectation and the standard deviation of the number of arrivals in the first 3 hours
2. Find the probability that at least one person arrives between 10:30 and 11:00.
3. Find the probability that exactly 3 people arrive between 5:00 and 7:00.
4. Given that 2 people arrived in the first hour, find the probability that 10 people arrive in the first three hours.

   Note: To make sense of the phrase *the first three hours*, you can measure time duration from any fixed time point.

Poisson process - example

Let $N(t)$ be the number of arrivals in t hours.

1. $N(3) \sim Pois(3.5 \times 3) = Pois(10.5)$. $E[N(3)] = 10.5$ and $SD[N(3)] = \sqrt{10.5} \approx 3.24$.
2. If s and t denote 10:30 and 11:00 respectively, let the number of customers arriving between those times be $X = N(t) - N(s) \sim Pois(1.75)$. Thus $P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-1.75} = 0.8262$
3. We need to find $P(X = 3)$ where $X \sim Pois(7)$, which is 0.0521.
4. Because of the independence between arrivals in the first hour and arrivals in the next two hours, the probability that 10 people arrive in the first three hours given that 2 people arrived in the first hour is the same as the probability of 8 people arriving in two hours. The number of arrivals in two hours has $Pois(7)$ distribution, so the required probability is $\frac{e^{-7}7^8}{8!} = 0.1304$.