5 The Binomial and Poisson Distributions

5.1 The Binomial distribution

- Consider the following circumstances (binomial scenario):
  1. There are \( n \) trials.
  2. The trials are independent.
  3. On each trial, only two things can happen.
     - We refer to these two events as success and failure.
  4. The probability of success is the same on each trial.
     - This probability is usually called \( p \).
  5. We count the total number of successes.
     - This is a discrete random variable, which we denote by \( X \), and which can take any value between 0 and \( n \) (inclusive).

- The random variable \( X \) is said to have a binomial distribution with parameters \( n \) and \( p \); abbreviated
  \[ X \sim \text{Bin}(n, p) \]

- It is easy to show that if \( X \sim \text{Bin}(n, p) \) then
  \[ P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k} \]
  for \( k = 0, 1, \ldots, n \).

- \( \binom{n}{k} \) is the binomial coefficient and is the number of sequences of length \( n \) containing \( k \) successes.
  \[ \binom{n}{k} = \frac{n!}{(n-k)!k!} \]

- The expectation and variance of \( X \) are given by
  \[ E[X] = np \]
  \[ \text{Var}[X] = np(1 - p) \]

The Binomial Distribution: Example

The shape of the distribution depends on \( n \) and \( p \).

![Binomial Distributions](image-url)
Example:
Suppose that it is known that 40% of voters support the Conservative party. We take a random sample of 6 voters. Let the random variable $Y$ represent the number in the sample who support the Conservative party.

1. Explain why the distribution of $Y$ might be binomial.
2. Write down the probability distribution of $Y$ as a table of probabilities.
3. Find the mean and variance of $Y$ directly from the probability distribution.


5.2 The Poisson distribution

- The binomial distribution is about counting successes in a fixed number of well-defined trials, i.e., $n$ is known and fixed.
- This can be limiting as many counts in science are open-ended counts of unknown numbers of events in time or space.
- Consider the following circumstances:
  1. Events occur randomly in time (or space) at a fixed rate $\lambda$
  2. Events occur independently of the time (or location) since the last event.
  3. We count the total number of events that occur in a time period $s$, and we let $X$ denote the event count.
- The random variable $X$ has a Poisson distribution with parameter $(\lambda s)$; abbreviated $X \sim \text{Po}(\lambda s)$
- If $X \sim \text{Po}(\lambda s)$ then $P[X = x] = e^{-\lambda s} \frac{(\lambda s)^x}{x!}$ for $k = 0, 1, 2, \ldots$.
- The expectation and variance of $X$ are given by $E[X] = \lambda s$ and $\text{Var}[X] = \lambda s$.

The Poisson Distribution
Like the binomial distribution, the shape of the Poisson distribution changes as we change its parameter.
Example: Yeast
Gossett, the head of quality control at Guiness brewery c. 1920 (and discoverer of the \( t \) distribution), arranged for counts of yeast cells to be made in sample vessels of fluid. He found that at a certain stage of brewing the counts were \( \text{Po}(0.6) \). Let \( X \) be the count from a sample. Find \( P[X \leq 3] \).

5.3 The Poisson approximation to the Binomial

The Poisson approximation to the Binomial

- Consider the Poisson scenario with events occurring randomly over a time period \( s \) at a fixed rate \( \lambda \).
- Now, split the time interval \( s \) into \( n \) subintervals of length \( s/n \) (very small).
- Let's consider each mini-interval as a “success” if there is an event in it.
- Now we have \( n \) independent trials with \( p \approx \frac{\lambda s}{n} \)
- The counts \( X \) are then binomial.
- If we assume there is no possibility of obtaining two events in the same interval, then we can say
  \[
  P[X = x] \approx P[T = x] = \binom{n}{x} \left( \frac{\lambda s}{n} \right)^x \left( 1 - \frac{\lambda s}{n} \right)^{n-x}
  \]
- It can be shown that as \( n \) increases and \( p \) decreases, this formula converges to
  \[
  e^{-\lambda s} \frac{(\lambda s)^x}{x!}
  \]
- Hence the Binomial distribution \( T \sim \text{Bin}(n, p) \), can be approximated by the Poisson \( T \sim \text{Po}(np) \) when \( np \) is small.
- This approximation is good if \( n \geq 20 \) and \( p \leq 0.05 \), and excellent if \( n \geq 100 \) and \( np \leq 10 \).

Example: Computer Chip Failure
A manufacturer claims that a newly-designed computer chip is has a 1% chance of failure because of overheating. To test their claim, a sample of 120 chips are tested. What is the probability that at least two chips fail on testing?

Suggested Exercises: Q30–34.

6 The Normal Distribution

6.1 The Normal Distribution

The Normal Distribution

- The most widely useful continuous distribution is the Normal (or Gaussian) distribution.
- In practice, many measured variables may be assumed to be approximately normal.
- Derived quantities such as sample means and totals can also be shown to be approximately normal.
- A rv \( X \) is Normal with parameters \( \mu \) and \( \sigma^2 \), written \( X \sim \text{N}(\mu, \sigma^2) \), when it has density function
  \[
  f_X(x) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left[ -\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 \right]
  \]
  for all real \( x \), and \( \sigma > 0 \).
The Normal Distribution

The Standard Normal

- The standard Normal random variable is a normal rv with \( \mu = 0 \), and \( \sigma^2 = 1 \). It is usually denoted \( Z \), so that \( Z \sim N(0, 1) \).
- The cumulative distribution function for \( Z \) is denoted \( \Phi(z) \) and is

\[
\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-1}{2}x^2\right) dx.
\]

- Unfortunately, there is no neat expression for \( \Phi(z) \), so in practice we must rely on tables (or computers) to calculate probabilities.

Properties of the Standard Normal & Tables

- \( \Phi(0) = 0.5 \) due to the symmetry
- \( P[a \leq Z \leq b] = \Phi(b) - \Phi(a) \).
- \( P[Z < -a] = \Phi(-a) = 1 - \Phi(a) = P[Z > a] \), for \( a \geq 0 \) – hence tables only contain probabilities for positive \( z \).
- \( \Phi \) is very close to 1 (0) for \( z > 3 \) (\( z < -3 \)) – most tables stop after this point.

Example

i Find the probability that a standard Normal rv is less than 1.6.

ii Find a value \( c \) such that \( P(-c \leq Z \leq c) = 0.95 \).

6.2 Standardisation

- If \( X \sim N(\mu, \sigma^2) \), then \( Z = \frac{X - \mu}{\sigma} \) is the standardized version of \( X \), and \( Z \sim N(0, 1) \).
- Even more importantly, the distribution function for any normal rv \( X \) is given by

\[
F_X(x) = \Phi\left(\frac{x - \mu}{\sigma}\right),
\]

and so the cumulative probabilities for any normal rv \( X \) can be expressed as probabilities of the standard normal \( Z \).
- This is why only the standard Normal distribution is tabulated.
Example

1. Let $X$ be $N(12, 25)$. Find $P[X > 3]$
2. Let $Y$ be $N(1, 4)$. Find $P[-1 < X < 2]$.

6.3 Other properties

Other properties

- Expectation and variance of $Z$:

\[
E[Z] = \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = 0, \quad \text{(integrand is an odd fn)}
\]
\[
E[Z^2] = 1, \quad \text{(integrate by parts)}
\]
\[
Var[Z] = 1.
\]

- Using our scaling properties it follows that for $X \sim N(\mu, \sigma^2)$,

\[
E[X] = \mu,
\]
\[
Var[X] = \sigma^2.
\]

- If $X$ and $Y$ are Normally distributed then the sum $S = X + Y$ is also Normally distributed (regardless of whether $X$ and $Y$ are independent).

6.4 Interpolation

Interpolation

- Normal distribution tables are limited and only give us values of $\Phi(Z)$ for a fixed number of $Z$.
- Often, we want to know $\Phi(Z)$ for values of $Z$ in between those listed in the tables.
- To do this we use linear interpolation - suppose we are interested in $\Phi(b)$, where $b \in [a, c]$ and we know $\Phi(a)$ and $\Phi(c)$.
- If we draw a straight line connecting $\Phi(a)$ and $\Phi(c)$ then (since $\Phi$ is smooth) we would expect $\Phi(b)$ to lie close to that line. Then

\[
\Phi(b) \simeq \Phi(a) + \left( \frac{b - a}{c - a} \right) (\Phi(c) - \Phi(a))
\]

Example

- Estimate the value of $\Phi(0.53)$ by interpolating between $\Phi(0.5)$ and $\Phi(0.6)$. 

6.5 Normal Approximation to the Binomial

- Regardless of \( p \), the \( \text{Bin}(n,p) \) histogram approaches the shape of the normal distribution as \( n \) increases. (This is actually a consequence of the strong law of large numbers; without going into more detail, the strong law simply says that certain distributions, under certain circumstances, converge to the normal distribution.)

- We can approximate the binomial distribution by a Normal distribution with the *same mean and variance*:
  
  \[
  \text{Bin}(n,p) \text{ is approximately } N(np,np(1-p))
  \]

- The approximation is acceptable when
  
  \[
  np \geq 10 \text{ and } n(1-p) \geq 10
  \]

  and the larger these values the better.

- For smallish \( n \), a *continuity correction* might be appropriate to improve the approximation.

- If \( X \sim \text{Bin}(n,p) \) and \( X' \sim N(np,np(1-p)) \), then
  
  \[
  P(X \leq k) \approx P(X' \leq k + 1/2)
  
  P(k_1 \leq X \leq k_2) \approx P(k_1 - 1/2 \leq X' \leq k_2 + 1/2)
  \]

**Example: Memory chips**

Let \( X_1, X_2, \) and \( X_3 \) be independent lifetimes of memory chips. Suppose that each \( X_i \) has a normal distribution with mean 300 hours and standard deviation 10 hours. Compute the probability that at least one of the three chips lasts at least 290 hours.

**Suggested Exercises**: Q35–38.