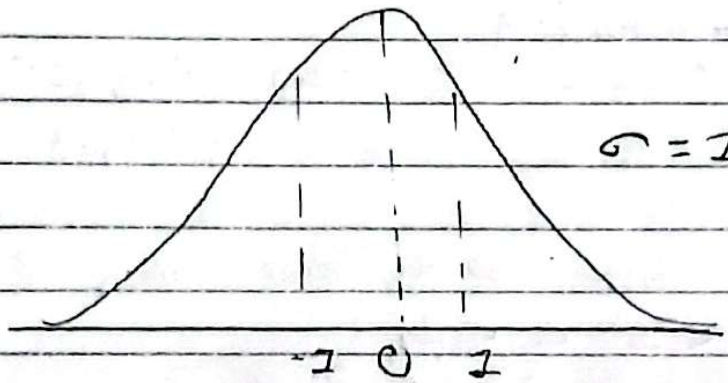
Summary of standard Normal Random Variable and Probability Computation

Definition

A standard normal random variable (Z) is a normally distributed random variable with Mean $\mu = 0$, standard deviation $\sigma = 1$

The density function for Z is illustrated in figure



Computing probabilities using Cumulative Tables

Instead of using density function, probabilities are computed from cumulative probability table which provide values of $P(Z < z)$

Example of Probability Computation

a $P(Z < 1.48)$

From the table $P(Z < 1.48) = 0.9306$

b $P(Z < -0.25)$

From the table $P(Z < -0.25) = 0.4013$

Finding $P(Z > z)$ using Complements

$P(Z > 1.60)$

Complement Rule: $P(Z > 1.60) = 1 - P(Z < 1.60)$

$= 1 - 0.9452$

$= 0.0548$

Finding $P(a < Z < b)$ for finite Intervals

$P(0.5 < Z < 1.57)$

$P(Z < 1.57) = 0.9418$

$P(Z < 0.50) = 0.6915$

Compute probability

$P(0.5 < Z < 1.57) = 0.9418 - 0.6915$

$= 0.2503$

Finding Computing Probabilities when Z is not in table

$P(1.13 < Z < 4.16)$

$P(Z < 1.13) = 0.8708$,

but $P(Z < 4.16)$ is not in the table (rounds to 1.000

Compute Probability

$P(1.13 < Z < 4.16) = 1.0000 - 0.8708$

$= 0.1292$

Empirical Rule for Normal Distribution

(a) $P(-1 < z < 1)$

from the table

$P(-1 < z < 1) = 0.8413 - 0.1587 = 0.683$

Interpretation: About 68% of the data falls within one standard deviation of the mean

(b) $P(-2 < z < 2)$

From the table

$P(-2 < z < 2) = 0.9772 - 0.0228 = 0.9544$

Interpretation: About 95% of the data falls within two standard deviations of the mean

(c) $P(-3 < z < 3)$

from the table

$P(-3 < z < 3) = 0.9987 - 0.0013 = 0.9974$

Interpretation: About 99.7% of data falls within three standard deviations of the mean

Key Takeaways

1 A standard normal random variable Z has $\mu = 0$ and $\sigma = 1$

2 Probabilities are obtained using cumulative normal probability tables

3 Empirical Rule

68% of values lie within $\pm 1\sigma$

95% of values lie within $\pm 2\sigma$

99.7% of values lie within $\pm 3\sigma$

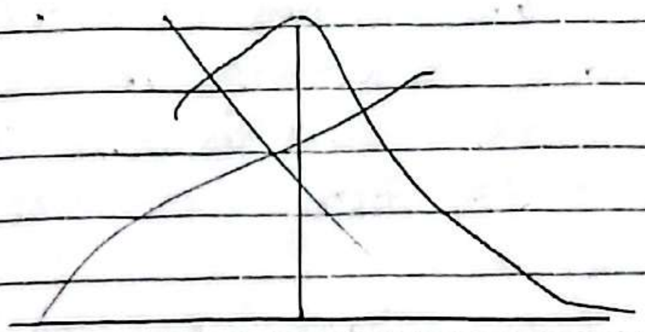
Summary: Normal Approximation to the Binomial Distribution
Binomial distribution and Probability Computation
- The binomial probability formula calculates the probability of obtaining x heads in N flips of a fair coin.

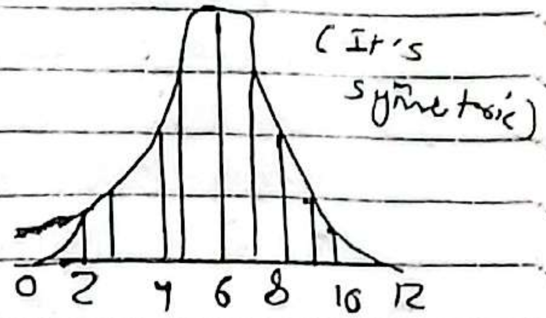
$$P(x) = \frac{N!}{x!(N-x)!}$$

Example finding the probability of getting 60 or more heads in 100 flips
- Recompute P(60), P(61), P(62), etc and Sum them
- Before calculators and computers, this process were extremely tedious.

de Moivre and the Normal Approximation
- Abraham de Moivre, an 18th-century statistician, observed that as N increased, the binomial distribution's shape approached a smooth curve
- He derived a mathematical expression for this curve, now called the normal distribution (or normal curve)

The Normal curves as an approximation.
- The normal curve closely approximates the binomial distribution when N is large.
- Figure demonstrates how well the normal distribution fits the binomial probabilities for 12 flips.





Importance of the Normal Distribution
- Many natural phenomena follow a normal distribution
- First application: Measurement errors in astronomy
  - Galileo observed that errors were symmetric, with small errors more common than large errors
  - In the 19th century, mathematicians Adrain (1808) and Gauss (1809) formalize the normal distribution formula for modeling errors

Central limit theorem and Normality
- Laplace (1778) discovered that sample means follow a normal distribution even if the original data is not normally distributed.
- This forms the basis of the Central limit Theorem, a key concept in Probability and statistics.

Applications of the Normal distribution
- Used in hypothesis testing, since mean differences are assumed to be normally distributed.
- Quetelet applied it to human characteristics (eg height, weight, strength)

Summary areas under the Normal distribution

Computing Areas under the Normal curve

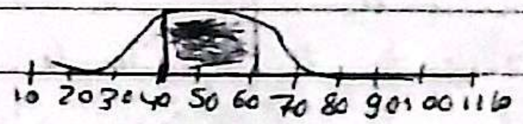
- The area under a normal distribution can be computed using calculus, but for practical purposes we use table and computer programs.
- The area represents probabilities and follows a consist pattern based on empirical Rule

Empirical Rule for Normal distributions.

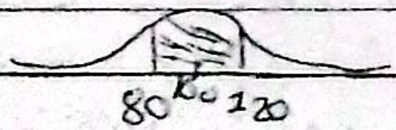
- 68% of the area lies within one standard deviation of the mean
- 95% of the area lies within 1.96 standard deviation of the mean
-

Examples of Normal Distributions

1 Mean = 50 SD = 10 ∴ 68% of the area is between 40 and 60



2 Mean = 100 SD = 20 ∴ 68% of the area is between 80 and 120



3 Mean = 75 SD = 10 95% of the area is between 55.4 and 94.6

Using a Normal Calculator

A normal calculator can compute the probability of values falling above, below, or between certain points in a normal distribution eg Mean: 90, SD = 12, $P(X > 110) = 0.0478$

Finding Percentiles Using the Inverse Normal Calculator

To find the 75th Percentile for a normal distribution with mean = 90, SD = 12

The score corresponding to the 75th percentile is 98.09

Normal approximation to the Binomial Distribution

Why use the Normal Approximation?

- The binomial distribution is discrete, while the normal distribution is continuous.
- The normal distribution provides a good approximation for the binomial distribution when N (number of trials) is large and π (probability of success) is not too close to 0 or 1.

Example: 6 Heads in 10 flips of a fair coin.

Given;

$N = 10, \pi = 0.5$

Mean $\mu = N\pi = 10 \times 0.5 = 5$

variance $\sigma^2 = N\pi(1-\pi)$ $5(0.5) = 2.5$

Standard deviation $\sigma = \sqrt{2.5} = 1.5811$

Continuity Correction

Since the normal distribution is continuous and binomal is discrete, we apply a continuity correction by considering the range 7.5 to 8.5 instead of just 8

Using the Standard Normal Distribution

1) Find Z-score for 8.5

$$Z = \frac{8.5 - 5}{1.5811} = 2.21$$

Area below $Z = 2.21$ = 0.987

2) Compute Approximate probability

P (8 heads) = 0.987 − 0.943 (Z score for 7.5)

= 0.044

This is a close approximation to the true binomial probability

Extending to a Range of values

To find $P(8 \leq X \leq 10)$, compute:

Area below 10.5 minus Area below 7.5

When is the Approximation Good?

- The normal approximation is accurate if:

$N\pi > 10$ and $N(1-\pi) > 10$

- If both values are greater than 10, the normal approximation works well.