Freebooks.pk

Free Notes • Past Papers • Guides for Students

ICS Part-II Statistics — Punjab Board

CHAPTER 10

Normal Distribution

Statistics • 2nd Year • ICS • Punjab Board • Free PDF Notes

Chapter 10: Normal Distribution

The normal distribution is the single most important probability distribution in all of statistics. It shows up everywhere -- heights and weights of individuals, IQ scores, measurement errors -- because it is also the limiting form that many other probability distributions approach, and it underlies the central limit theorem that makes so much of inferential statistics possible. This chapter opens the second half of the ICS Statistics course: everything from Chapter 11 onward (sampling distributions, estimation, hypothesis testing) leans on the normal distribution introduced here.

The chapter builds the normal probability density function and its key properties, then introduces the standard normal random variable Z -- a single reference distribution (mean 0, variance 1) that lets us answer probability questions for ANY normal distribution using one shared set of tables, through the simple transformation Z = (X - mu)/sigma.

Learning Objectives

- State the normal probability density function and identify its two parameters, mu and sigma
- List the key properties of the normal distribution: symmetry, mode, special areas, moments
- Standardize any normal random variable X into the standard normal variable Z
- Use the standard normal cumulative distribution function Phi(z) to calculate probabilities
- Find quantiles and percentiles of the standard normal distribution and de-standardize them back to X
- Solve applied problems involving normally distributed variables: scores, heights, dimensions, times
- Work backward to find mu or sigma (or both) given probability information

Key Concepts

10.1 The Normal Probability Density Function

A continuous random variable X is normally distributed if its probability density function is f(x) = [1/(sigma.sqrt(2.pi))] . e^[-(1/2)((x-mu)/sigma)^2], for -infinity < x < infinity, where mu can be any real number and sigma must be positive. A normal distribution is fully characterized by just these two parameters -- mu (its mean) and sigma (its standard deviation) -- and is commonly written X ~ N(mu, sigma-squared).

The graph of this function is called a normal curve: unimodal, symmetrical, and bell-shaped. The parameter sigma controls how flat or peaked the curve is -- decreasing sigma (with mu fixed) makes the curve more sharply peaked (higher probability of X being close to mu); increasing sigma flattens it. Changing mu (with sigma fixed) simply slides the same-shaped curve left or right along the x-axis.

10.1 Properties of the Normal Distribution

The normal distribution has several defining properties worth memorizing. It is a continuous distribution ranging from -infinity to +infinity, with total area under the curve equal to 1 (P(-infinity < X < infinity) = 1). It is unimodal with its single peak (mode) at x = mu, where the maximum ordinate is f(mu) = 1/(sigma.sqrt(2.pi)). It is perfectly symmetrical, so mean = median = mode = mu, the lower and upper quartiles are equidistant from mu, and all odd-order moments about the mean are zero.

Three 'special areas' recur constantly in problems: P(mu-sigma < X < mu+sigma) = 0.6827 (about 68%), P(mu-2sigma < X < mu+2sigma) = 0.9545 (about 95%), and P(mu-3sigma < X < mu+3sigma) = 0.9973 (about 99.7%) -- together known as the empirical rule. The quartile deviation is Q.D.(X) = 0.6745.sigma, the mean deviation is M.D.(X) = 0.7979.sigma, and the variance/SD are simply sigma-squared and sigma. The curve is asymptotic to the x-axis (never touches it) and has points of inflexion at x = mu - sigma and x = mu + sigma. A useful reproductive property: if X1 ~ N(mu1, sigma1^2) and X2 ~

N(mu2, sigma2^2) are independent, then X1+X2 ~ N(mu1+mu2, sigma1^2+sigma2^2).

10.2 The Standard Normal Random Variable

Because the normal cumulative distribution function cannot be integrated in closed form, and because there are infinitely many possible (mu, sigma) combinations, statisticians instead standardize: any normal random variable X can be transformed into the standard normal random variable Z = (X-mu)/sigma, which always has mean 0 and variance 1, regardless of X's original mu and sigma. This lets a single set of tables answer probability questions for every normal distribution.

The standard normal density function is written phi(z) = [1/sqrt(2.pi)].e^(-z^2/2), and its cumulative distribution function is written Phi(z) = P(Z <= z) -- the area under the standard normal curve up to z. Because the curve is symmetric, phi(-z) = phi(z). Four handy results follow directly from the definition: P(Z<=a) = Phi(a); P(Z>=a) = 1-Phi(a); P(a<=Z<=b) = Phi(b)-Phi(a); and using symmetry, Phi(-a) = 1-Phi(a), so P(|Z|<a) = 2.Phi(a)-1 and P(|Z|>a) = 2.Phi(-a). Phi(z) and its inverse Phi-inverse(p) are both read from standard tables.

10.2 Quantiles, Percentiles, and De-standardizing

The p-th quantile (or 100p-th percentile) of Z is the value z_p such that Phi(z_p) = p, i.e., z_p = Phi-inverse(p) -- found by looking up p in the cumulative table and reading off the corresponding z. Once a quantile z_p of the standard normal distribution is known, it can be de-standardized back into the original scale of any normal variable X using x_p = mu + sigma.z_p, obtained by rearranging Z = (X-mu)/sigma into X = mu + sigma.Z.

For applied problems on any normal variable X ~ N(mu, sigma-squared), the same standardization logic applies directly to probability statements: P(X<=a) = Phi[(a-mu)/sigma]; P(X>=a) = 1-Phi[(a-mu)/sigma]; P(a<=X<=b) = Phi[(b-mu)/sigma] - Phi[(a-mu)/sigma]. This single technique -- standardize, look up in the Z table, and de-standardize when needed -- solves essentially every problem type in this chapter, from finding probabilities given X-values to finding X-values (or even mu or sigma) given probabilities.

Important Definitions

What is a normal distribution?: A continuous probability distribution described by f(x) = [1/(sigma.sqrt(2.pi))].e^[-(1/2)((x-mu)/sigma)^2], fully determined by its mean mu and standard deviation sigma, producing a symmetric, unimodal, bell-shaped curve.

What is the standard normal distribution?: The special case of the normal distribution with mean 0 and variance 1, obtained from any normal variable X via the transformation Z = (X-mu)/sigma.

What is Phi(z)?: The standard normal cumulative distribution function: Phi(z) = P(Z<=z), the area under the standard normal curve up to the point z.

What is phi(z)?: The standard normal probability density function (ordinate): phi(z) = [1/sqrt(2.pi)].e^(-z^2/2), the height of the standard normal curve at the point z.

What is a quantile (percentile) of the standard normal distribution?: The value z_p such that P(Z<=z_p) = p; found as z_p = Phi-inverse(p) using the inverse standard normal table.

What does 'de-standardizing' mean?: Converting a standard normal value z back to the original scale of X using X = mu + sigma.Z.

What is the empirical rule for the normal distribution?: Approximately 68.27% of values lie within 1 standard deviation of the mean, 95.45% within 2 standard deviations, and 99.73% within 3 standard deviations.

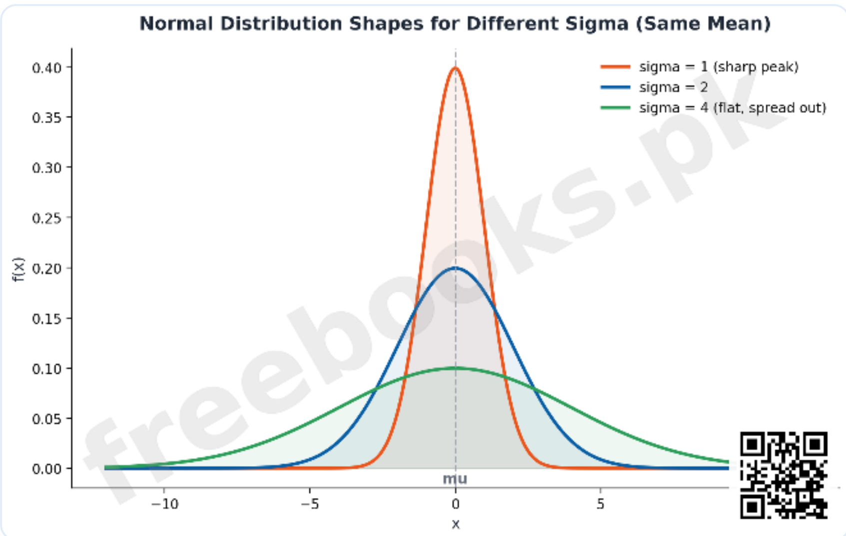
Key Facts and Relations

Topic	Key Fact / Relation
Normal PDF	f(x) = [1/(sigma.sqrt(2.pi))] e^[-(1/2)((x-mu)/sigma)^2]
Standardization	Z = (X - mu) / sigma
De-standardizing	X = mu + sigma.Z
Standard normal PDF	phi(z) = [1/sqrt(2.pi)] e^(-z^2/2)

Topic	Key Fact / Relation
Standard normal CDF	Phi(z) = P(Z <= z)
Symmetry results	Phi(-a) = 1-Phi(a); P(|Z|<a) = 2Phi(a)-1; P(|Z|>a) = 2Phi(-a)
Quartile deviation	Q.D.(X) = 0.6745 sigma
Mean deviation	M.D.(X) = 0.7979 sigma
Special areas (empirical rule)	mu+/-sigma: 68.27% | mu+/-2sigma: 95.45% | mu+/-3sigma: 99.73%
Quantile of X	x_p = mu + sigma . z_p, where z_p = Phi-inverse(p)
Moments about mean	mu1=0, mu2=sigma^2, mu3=0, mu4=3sigma^4; Beta1=0, Beta2=3

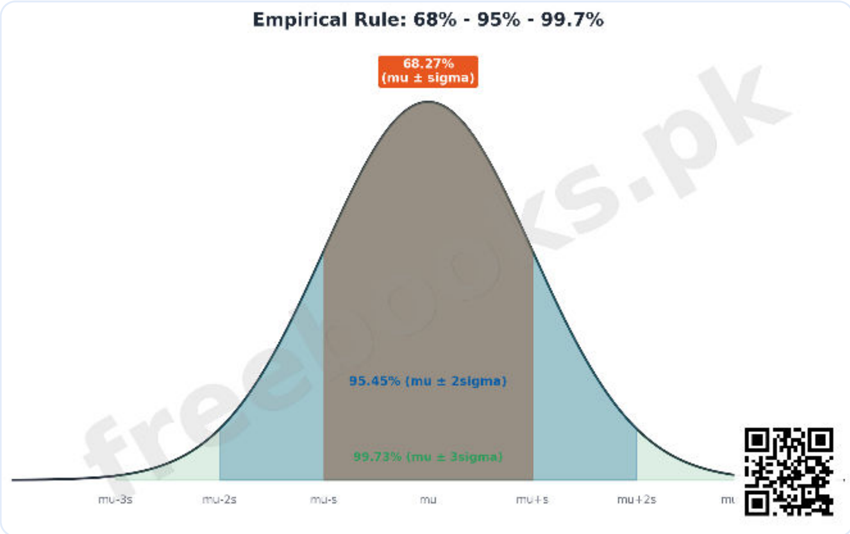
Diagrams

Normal Distribution Shapes for Different Sigma: Three normal curves with the same mean but different standard deviations, showing how a smaller sigma produces a sharper, more peaked curve while a larger sigma produces a flatter, more spread-out curve



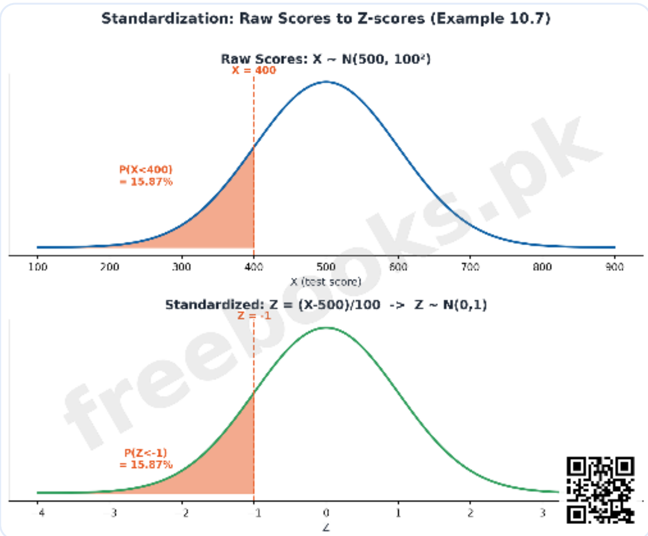
Normal Distribution Shapes for Different Sigma

Empirical Rule: 68-95-99.7: The standard normal curve with the areas within 1, 2, and 3 standard deviations of the mean shaded and labelled 68.27%, 95.45%, 99.73%, illustrating the special areas under the normal curve



Empirical Rule: 68-95-99.7

Standardization: Raw Scores to Z-scores: Two aligned normal curves -- the raw test-score distribution (mean 500, SD 100) on top and the standard normal Z distribution below -- with the region X<400 (Z<-1) shaded on both, illustrating the standardization technique from Example 10.7



Standardization: Raw Scores to Z-scores

Short Questions & Answers

What are the two parameters of a normal distribution and what do they control?

Mu (the mean) controls the location/center of the curve; sigma (the standard deviation) controls its spread -- smaller sigma gives a sharper peak, larger sigma gives a flatter curve.

Why do we standardize a normal random variable before using probability tables?

Because there are infinitely many possible normal distributions (one for every combination of mu and sigma), so probabilities are tabulated only once for the single standard normal distribution (mean 0, variance 1), and any normal variable is converted to this common scale via Z=(X-mu)/sigma.

State the empirical rule (68-95-99.7 rule) for a normal distribution.

About 68.27% of values fall within 1 standard deviation of the mean, about 95.45% within 2 standard deviations, and about 99.73% within 3 standard deviations.

Why is the mode of a normal distribution equal to its mean?

Because the normal distribution is perfectly symmetric and unimodal, its single peak occurs exactly at x=mu, and by symmetry the mean, median, and mode all coincide at that same point.

What is the maximum ordinate of a standard normal curve, and where does it occur?

The maximum ordinate is phi(0) = 1/sqrt(2.pi) ≈ 0.3989, occurring at z=0.

How do you find the value of X corresponding to a given percentile of a normal distribution?

First find the corresponding z-value using the standard normal table (z_p = Phi-inverse(p)), then de-standardize using x_p = mu + sigma.z_p.

Long Questions & Answers

Explain the normal probability density function and its key properties, and explain why the normal distribution is considered the most important distribution in statistics.

The normal distribution occupies a uniquely central position in statistical theory and practice, and understanding why requires looking both at its mathematical structure and at the practical reasons it appears so often in real data. Formally, a continuous random variable X is said to be normally distributed if its probability density function takes the specific form f(x) equals one over sigma times the square root of 2 pi, multiplied by e raised to the power of negative one-half times the quantity (x minus mu) divided by sigma, all squared, for x ranging across the entire real number line from negative infinity to positive infinity. This formula might look intimidating at first glance, but its behavior is governed entirely by just two parameters: mu, which can be any real number and represents the mean of the distribution, and sigma, which must always be strictly positive and represents the standard deviation. Because exactly these two numbers completely determine the shape and location of the entire curve, statisticians commonly abbreviate the statement 'X follows a normal distribution with mean mu and variance sigma-squared' using the compact notation X tilde N(mu, sigma-squared). The graph traced out by this density function is called the normal curve, and it possesses several visually and mathematically striking characteristics: it is unimodal, meaning it has exactly one peak; it is perfectly symmetrical around that peak; and it forms the instantly recognizable bell shape that gives the distribution its popular nickname. The specific role played by each of the two parameters becomes clear once we consider what happens as each is varied independently of the other. Holding mu fixed while decreasing sigma causes the curve to become more sharply peaked and narrow, concentrating probability more tightly around the mean and thereby making values close to mu considerably more likely relative to values far away from it. Conversely, holding mu fixed while increasing sigma causes the curve to flatten out and spread wider, distributing probability more evenly across a broader range and correspondingly making values close to mu comparatively less likely. Meanwhile, if sigma is instead held constant while mu itself is varied, the overall shape of the curve remains completely unchanged, and only its horizontal position along the number line shifts, sliding to center on whatever new value mu has taken. Beyond this basic shape-and-location behavior, the normal distribution obeys several formal properties that prove immensely useful when solving problems. Because the entire distribution is a valid continuous probability distribution, the total area beneath the curve across its entire infinite range must equal exactly 1, and because probability under any continuous distribution corresponds to area rather than height at a single point, the probability of X falling in any particular interval is found by calculating the corresponding area under the curve above that interval. Because of the distribution's perfect symmetry, several familiar summary measures all collapse onto the exact same single value: the mean, the median, and the mode of any normal distribution are always identical, all located precisely at mu, and this same symmetry also guarantees that the lower and upper quartiles sit at exactly equal distances on either side of mu, and that every odd-numbered moment calculated about the mean equals exactly zero. Perhaps the most practically useful properties of all are the so-called special areas that hold true for absolutely every normal distribution regardless of its specific mu and sigma values: no matter which particular normal distribution is being considered, exactly 68.27 percent of its total probability always falls within one standard deviation of the mean, 95.45 percent always falls within two standard deviations, and 99.73 percent always falls within three standard deviations -- a remarkably consistent pattern collectively known as the empirical rule, which turns out to be enormously useful for quickly estimating probabilities without needing to consult any table at all. As for why this particular distribution earns its status as the single most important distribution across the whole of statistics, three distinct but mutually reinforcing reasons stand out clearly. First, purely as an empirical matter of observation, an enormous number of naturally occurring variables genuinely do follow something extremely close to a normal distribution in real life, including the heights and weights of individuals within a population, standardized intelligence test scores, and the small random errors that inevitably accompany essentially any careful physical measurement. Second, on a more theoretical level, the normal distribution frequently emerges mathematically as the limiting form that numerous entirely different probability distributions converge toward under suitable conditions, which means it can often serve as a remarkably accurate practical approximation to those other distributions even in situations where the underlying process generating the data isn't literally normal to begin with. Third, and perhaps most profoundly important of all for the rest of statistical theory, the normal distribution is also precisely the limiting distribution described by the celebrated central limit theorem, a foundational and far-reaching result explored in depth in the following chapter, which explains why sample means calculated from repeated random sampling tend to follow an approximately normal distribution even when the underlying population itself is not normally distributed at all -- and it is this specific theoretical property above all others that ultimately underlies the vast majority of statistical inference techniques used throughout the entire discipline, from constructing confidence intervals to performing hypothesis tests.

Explain the concept of the standard normal random variable, describe the process of standardization and de-standardization, and explain how the standard normal table is used to solve problems involving probabilities, quantiles, and unknown parameters.

Since a genuinely unlimited number of different normal distributions exist mathematically, one for literally every possible combination of a mean mu and a standard deviation sigma, it would be entirely impractical, in fact essentially impossible, to construct and maintain a separate complete probability table for every single one of these countless distributions individually. Statisticians resolve this fundamental practical difficulty through an elegant technique called standardization, which converts any and every normal random variable into one single common, universally shared reference distribution before any table lookup is ever performed. Specifically, given any normal random variable X that follows a distribution with mean mu and standard deviation sigma, the standardized variable Z is defined directly by the simple transformation formula Z equals the quantity (X minus mu), all divided by sigma. What makes this particular transformation so remarkably powerful and useful in practice is a clean mathematical guarantee: no matter what specific numerical values the original mu and sigma of X happened to be, the resulting standardized variable Z that comes out the other end of this transformation will always, without any exception whatsoever, have a mean of exactly 0 and a variance of exactly 1. This special, universally shared distribution is formally called the standard normal distribution, conventionally written as Z tilde N(0,1), and because literally every normal distribution can be converted into this exact same standard form through the identical transformation process, one single unified probability table can then be constructed just once, published, and subsequently reused again and again to solve probability problems for every conceivable normal distribution, regardless of that particular distribution's own specific mu and sigma values. The standard normal distribution comes equipped with its own dedicated and specially reserved mathematical notation precisely because it gets referenced so extremely frequently throughout the rest of statistical work. Its probability density function, representing the height or ordinate of the curve at any given point z, is written using the lowercase Greek letter phi, as phi(z) equals one over the square root of 2 pi, multiplied by e raised to the power of negative z-squared divided by 2. Its cumulative distribution function, representing instead the total accumulated area under that same curve running from negative infinity all the way up to some specified point z, is written using the corresponding uppercase Greek letter, as Phi(z) equals P(Z is less than or equal to z). This particular uppercase Phi(z) function is unquestionably the single most frequently used and referenced tool throughout this entire chapter, since it directly supplies the answer to the single most fundamental and recurring question posed again and again in problems of this type: what proportion, or equivalently what probability, of a standard normal distribution's total area lies at or below some particular specified point along the z-axis. Because the underlying standard normal curve possesses perfect mirror symmetry around the central value of zero, several extremely useful and time-saving relationships follow immediately and directly as natural consequences of that inherent symmetry, all readily available without requiring any additional or separate calculation whatsoever: the probability that Z exceeds some point a can always be found quickly as simply 1 minus Phi(a); the probability that Z falls somewhere between two specified points a and b can always be found as the straightforward difference Phi(b) minus Phi(a); and by that same underlying symmetry, Phi of negative a always equals exactly 1 minus Phi(a), a relationship which itself then leads directly to two further convenient shortcut formulas frequently needed for absolute-value probability statements, specifically that P(the absolute value of Z is less than a) equals 2 times Phi(a) minus 1, and that P(the absolute value of Z is greater than a) equals 2 times Phi(negative a). Beyond simply calculating probabilities in the forward direction starting from some given z-value, problems in this chapter frequently also demand working the entire process in reverse: given some specific target probability or percentage value already in hand, the task instead becomes finding whichever particular z-value corresponds to producing exactly that stated probability. This particular kind of reverse lookup defines what is formally called a quantile, or equivalently and interchangeably a percentile, of the standard normal distribution: the p-th quantile, conventionally denoted using the symbol z sub p, is specifically defined as that unique value satisfying the defining equation Phi(z sub p) equals p, which is written more compactly and directly using standard inverse-function notation as z sub p equals Phi inverse of p, and this particular inverse value is found in practice simply by locating the target probability p somewhere within the body of the standard cumulative probability table and then reading directly across to whichever corresponding z-value produced that exact same tabulated probability in the first place. Once any particular quantile of the standard normal distribution Z has successfully been determined through this table-lookup process, that resulting standardized value can then, whenever it becomes genuinely necessary or useful for the problem at hand, be converted straight back once again into the original measurement scale belonging to whatever specific normal variable X was actually being studied in the first place, through a process referred to quite naturally and appropriately as de-standardizing. This reverse conversion is accomplished simply and directly by algebraically rearranging the original standardization formula Z equals (X minus mu) divided by sigma into its equivalent inverted form X equals mu plus sigma times Z, meaning that the p-th quantile expressed specifically in the original X-scale, conventionally written as x sub p, can always be calculated directly and immediately as x sub p equals mu plus sigma multiplied by z sub p. Taken together as a single unified, integrated procedure, this complete three-stage process, comprising standardizing the specific problem at hand into the common Z-scale, then consulting the single shared standard normal table to either find a probability or alternatively to find a percentile as the specific problem demands, and finally de-standardizing that resulting z-value back into the original problem's own particular scale whenever that final conversion step happens to be required, together forms the single foundational technique that runs consistently through virtually every single problem type encountered throughout this entire chapter -- whether the specific task at hand is calculating some probability directly from a given X-value, or instead working entirely in reverse to find some unknown X-value starting from a given target probability, or even solving the comparatively more advanced style of problem that requires determining an entirely unknown mean mu or unknown standard deviation sigma (or occasionally even solving simultaneously for both unknowns at once) by making skillful and deliberate use of whatever specific probability information happens to already be given as part of that particular problem's stated conditions.

Multiple Choice Questions (MCQs)

The normal distribution is fully characterized by which two parameters? (A) Median and mode (B) Mean (mu) and standard deviation (sigma) (C) Range and IQR (D) Skewness and kurtosis

Correct answer: (B) Mean (mu) and standard deviation (sigma). A normal distribution X ~ N(mu, sigma^2) is completely determined by its mean mu and standard deviation sigma.

The total area under any normal curve equals: (A) 0 (B) 0.5 (C) 1 (D) Infinity

Correct answer: (C) 1. Like all valid probability density functions, the total area under the normal curve is exactly 1.

In a normal distribution, mean, median, and mode: (A) Are always different (B) Are always equal (C) Sum to zero (D) Equal sigma

Correct answer: (B) Are always equal. Due to the perfect symmetry of the normal distribution, mean = median = mode = mu.

The standard normal distribution has mean and variance: (A) Mean 1, variance 0 (B) Mean 0, variance 1 (C) Mean mu, variance sigma^2 (D) Mean 100, variance 15

Correct answer: (B) Mean 0, variance 1. The standard normal distribution Z ~ N(0,1) always has mean 0 and variance 1.

The standardization formula is: (A) Z = X + mu/sigma (B) Z = (X - mu)/sigma (C) Z = X.sigma - mu (D) Z = mu/X

Correct answer: (B) Z = (X - mu)/sigma. Standardization converts any normal variable X into Z using Z = (X-mu)/sigma.

Approximately what percent of a normal distribution lies within 2 standard deviations of the mean? (A) 68.27% (B) 95.45% (C) 99.73% (D) 50%

Correct answer: (B) 95.45%. By the empirical rule, about 95.45% of values lie within mu +/- 2sigma.

Phi(z) represents: (A) The ordinate of the standard normal curve at z (B) P(Z <= z) (C) The variance of Z (D) The mean of X

Correct answer: (B) P(Z <= z). Phi(z) is the standard normal cumulative distribution function, giving P(Z<=z).

If P(Z<a) = 1 - P(Z<-a), this reflects which property? (A) Skewness (B) Kurtosis (C) Symmetry of the standard normal curve (D) The empirical rule

Correct answer: (C) Symmetry of the standard normal curve. This relationship, Phi(-a)=1-Phi(a), follows directly from the symmetry of the standard normal curve about zero.

The formula to de-standardize a z-value back to the original scale is: (A) X = mu - sigma.Z (B) X = mu + sigma.Z (C) X = sigma/mu (D) X = mu.Z

Correct answer: (B) X = mu + sigma.Z. De-standardizing uses X = mu + sigma.Z, the inverse of the standardization formula.

The maximum ordinate of the standard normal curve occurs at: (A) z = 1 (B) z = -1 (C) z = 0 (D) z = mu

Correct answer: (C) z = 0. The standard normal curve peaks at z=0, where phi(0) = 1/sqrt(2.pi) ≈ 0.3989.

Quick Revision Summary

- Normal PDF: f(x) = [1/(sigma.sqrt(2pi))] e^[-(1/2)((x-mu)/sigma)^2], parameters mu and sigma
- Symmetric, unimodal, bell-shaped; mean=median=mode=mu; total area = 1
- Empirical rule: mu+/-sigma = 68.27% | mu+/-2sigma = 95.45% | mu+/-3sigma = 99.73%
- Standardization: Z = (X-mu)/sigma; standard normal Z~N(0,1)
- Phi(z) = P(Z<=z); P(Z>a)=1-Phi(a); P(a<Z<b)=Phi(b)-Phi(a)
- Symmetry: Phi(-a)=1-Phi(a); P(|Z|<a)=2Phi(a)-1; P(|Z|>a)=2Phi(-a)

- Quantile: z_p = Phi-inverse(p); de-standardize: x_p = mu + sigma.z_p
- Q.D.(X) = 0.6745 sigma; M.D.(X) = 0.7979 sigma; Var(X)=sigma^2

Exam Tips

- Always sketch a quick bell curve and shade the region asked for — it prevents sign errors with Phi(a) vs 1-Phi(a)
- For 'P(X between a and b)' problems, standardize BOTH endpoints before subtracting Phi values
- For 'find X given a probability' problems, always look up the z-value FIRST, then de-standardize with x = mu + sigma.z
- Remember Phi(-z) = 1 - Phi(z) — most tables only list positive z, so use this to handle negative z-values
- When solving for unknown mu or sigma, set up the standardization equation and solve algebraically — don't guess
- Watch for continuity correction language ('less than 60mm' on discrete/ measured data may mean 'less than 59.5' in a problem's intended interpretation)