

Statistical Techniques for Assessment of Existing Concrete Structures—Report

Reported by ACI Innovation Task Group 11

ACI PRC-ITG-11-24



American Concrete Institute
Always Advancing



Documents

Document Sharing Hub



Statistical Techniques for Assessment of Existing Concrete Structures—Report

Copyright by the American Concrete Institute, Farmington Hills, MI. All rights reserved. This material may not be reproduced or copied, in whole or part, in any printed, mechanical, electronic, film, or other distribution and storage media, without the written consent of ACI.

The technical committees responsible for ACI committee reports and standards strive to avoid ambiguities, omissions, and errors in these documents. Despite these efforts, the users of ACI documents occasionally find information or requirements that may be subject to more than one interpretation or may be incomplete or incorrect. Users who have suggestions for the improvement of ACI documents are requested to contact ACI via the errata website at <http://concrete.org/Publications/DocumentErrata.aspx>. Proper use of this document includes periodically checking for errata for the most up-to-date revisions.

ACI committee documents are intended for the use of individuals who are competent to evaluate the significance and limitations of its content and recommendations and who will accept responsibility for the application of the material it contains. Individuals who use this publication in any way assume all risk and accept total responsibility for the application and use of this information.

All information in this publication is provided “as is” without warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose or non-infringement.

ACI and its members disclaim liability for damages of any kind, including any special, indirect, incidental, or consequential damages, including without limitation, lost revenues or lost profits, which may result from the use of this publication.

It is the responsibility of the user of this document to establish health and safety practices appropriate to the specific circumstances involved with its use. ACI does not make any representations regarding health and safety issues and the use of this document. The user must determine the applicability of all regulatory limitations before applying the document and must comply with all applicable laws and regulations, including but not limited to, United States Occupational Safety and Health Administration (OSHA) health and safety standards.

Participation by governmental representatives in the work of the American Concrete Institute and in the development of Institute standards does not constitute governmental endorsement of ACI or the standards that it develops.

ACI documents are written via a consensus-based process. The characteristics of ACI technical committee operations include:

- (a) Open committee membership
- (b) Balance/lack of dominance
- (c) Coordination and harmonization of information
- (d) Transparency of committee activities to public
- (e) Consideration of views and objections
- (f) Resolution through consensus process

The technical committee documents of the American Concrete Institute represent the consensus of the committee and ACI. Technical committee members are individuals who volunteer their services to ACI and specific technical committees.

38800 Country Club Drive
Farmington Hills, MI 48331
Phone: +1.248.848.3700
Fax: +1.248.848.3701

Statistical Techniques for Assessment of Existing Concrete Structures—Report

Report prepared by Innovation Task Group 11

Peter Barlow, Chair

Nicholas R. Triandafilou*, Secretary

Ali Abu-Yosef
F. Michael Bartlett*
Ziad Elaghoury

Jeremiah D. Fasl
Kip Gatto
Timothy R. W. Gillespie

Lawrence F. Kahn*
Ming Liu
John S. Lund

Diego Romero
Kyle D. Stanish*
Jeffrey F. West*

*Authors of this report.

This report demonstrates how statistical techniques can be used in the assessment of concrete structures to evaluate their condition and reliability and to assess the relative merits of possible repair options. A Reliability Toolkit provides licensed design professionals with a summary of some statistical methods that are useful for structural assessment. The report also gives two detailed examples of assessments for structural concrete repair.

Keywords: confidence level; data analysis; durability; material properties; mean load method; rehabilitation; reliability; sample size; testing.

ACI Committee Reports and Guides are intended for guidance in planning, designing, executing, and inspecting construction. This document is intended for the use of individuals who are competent to evaluate the significance and limitations of its content and recommendations and who will accept responsibility for the application of the information it contains. ACI disclaims any and all responsibility for the stated principles. The Institute shall not be liable for any loss or damage arising therefrom. Reference to this document shall not be made in contract documents. If items found in this document are desired by the Architect/Engineer to be a part of the contract documents, they shall be restated in mandatory language for incorporation by the Architect/Engineer.

CONTENTS

CHAPTER 1—INTRODUCTION AND SCOPE, p. 2

- 1.1—Introduction, p. 2
- 1.2—Scope, p. 2

CHAPTER 2—NOTATION, p. 2

CHAPTER 3—ASSESSMENT, p. 4

- 3.1—Planning, p. 4
- 3.2—Objective of testing program, p. 4
- 3.3—Properties and testing, p. 4

CHAPTER 4—BACKGROUND ON STATISTICAL TECHNIQUES, p. 5

- 4.1—Defining the population and determining sample sizes, p. 5
- 4.2—Characterizing data, p. 7
- 4.3—Comparing and combining datasets, p. 9
- 4.4—Bounding sample statistics using confidence intervals, p. 10
- 4.5—Determining fractile values, p. 13
- 4.6—Characterizing data representing extreme transient loads, p. 13

CHAPTER 5—RELIABILITY-BASED DURABILITY ASSESSMENT EXAMPLE, p. 14

- 5.1—Problem statement, p. 14

ACI PRC-ITG-11-24 was published August 2024.

Copyright © 2024, American Concrete Institute.

All rights reserved including rights of reproduction and use in any form or by any means, including the making of copies by any photo process, or by electronic or mechanical device, printed, written, or oral, or recording for sound or visual reproduction or for use in any knowledge or retrieval system or device. All rights reserved. This document is obtained from the copyright proprietors.



Documents

Document Sharing Hub



- 5.2—Structure details, p. 14
- 5.3—Assessment and testing methodology, p. 15
- 5.4—Condition assessment plan, p. 15
- 5.5—Reinforcement cover measurements, p. 16
- 5.6—Carbonation depth measurements, p. 18
- 5.7—Evaluation results and conclusion, p. 19

CHAPTER 6—RELIABILITY-BASED STRENGTH ASSESSMENT EXAMPLE, p. 20

- 6.1—Problem statement, p. 21
- 6.2—Preliminary assessment, p. 21
- 6.3—Detailed assessment: Case of good documentation, p. 22
- 6.4—Detailed assessment: Case of no documentation, p. 23
- 6.5—Strength assessment by mean load method, p. 27
- 6.6—Comparison of methods, p. 31
- 6.7—Summary, p. 31

CHAPTER 7—SUMMARY, p. 31

CHAPTER 8—REFERENCES, p. 32

Authored documents, p. 33

APPENDIX A—RELIABILITY TOOLKIT FOR ASSESSMENT OF EXISTING CONCRETE STRUCTURES, p. 33

- A.1—Introduction, p. 33
- A.2—Reliability index and mean load method, p. 33
- A.3—Target reliability indexes, p. 34
- A.4—Combinations of independent random variables, p. 34
- A.5—Statistical parameters for resistance, p. 35
- A.6—Extreme value distributions for transient load effects, p. 36
- A.7—Defining “population” for sampling, p. 37
- A.8—Data analysis, p. 37
- A.9—Bayes’ theorem, p. 37
- A.10—Examples, p. 38

CHAPTER 1—INTRODUCTION AND SCOPE

1.1—Introduction

Licensed design professionals working with existing structures are often challenged with developing an appropriate scope of proposed work during an assessment or investigation phase to characterize the as-built and current conditions of a structure. While there are statistical techniques that can be used to establish appropriate sample sizes, in general, there is no “one-size-fits-all” approach to assessing structural condition. Performance-based standards for assessment, repair, and rehabilitation of existing concrete structures, such as **ACI CODE-562**, require methods for representative measurement of such performance. As an example, **ACI CODE-562** uses statistical techniques for determining equivalent specified concrete compressive strength f_{ceq} for the strength assessment of existing structures. For other measurements needed for durability and structural assessments, testing and field investigation techniques are exten-

sively documented but there is little guidance on adoption of statistical methods to reliably characterize the necessary parameters. Information on the use of statistical methods for assessment and repair of existing concrete structures is needed because the conventional, well-known methods used in manufacturing industries for quality assurance are often not practical for single, existing structures.

This report demonstrates how statistical techniques can be used to assess the performance of existing concrete structures and of proposed repairs. A Reliability Toolkit, included as **Appendix A**, provides licensed design professionals with a summary of some statistical methods that are useful for structural assessment. The report also gives two detailed examples of assessments for structural concrete repair. The first example assesses durability, while the second evaluates strength. Each illustrates how methods from the Toolkit are used to quantify field-measured data to a level of confidence determined by the licensed design professional.

Developing an adequate assessment is important in terms of technical effectiveness and cost efficiency. **Chapter 3** provides guidance on which parameters should be quantified by testing, and **Chapter 4** discusses sample sizes and the statistical analysis of test results. The durability example presented in **Chapter 5** uses statistical analysis of cover and carbonation depth measurements to derive probable end-of-service-life limits. The strength example in **Chapter 6** illustrates structural assessment using preliminary, detailed, and reliability-based techniques specified in **ACI CODE-562**. The Reliability Toolkit in **Appendix A** may become the most frequently used part of this document as users become familiar with reliability-based methods.

The toolkit and examples include applications of functions available in the Microsoft Excel software; similar functions are available in other software packages and these example applications should not be viewed as an endorsement of a particular software package.

1.2—Scope

The scope of this report is limited to providing a summary of statistical techniques that are applicable to the assessment of existing concrete structures and to providing two detailed examples of such applications. Further, an introduction to reliability computations is presented as a Reliability Toolkit. The report is not intended to present all statistical methods or reliability concepts applicable to concrete design, construction, and quality assurance applications. Instead, it is a succinct primer targeted to practitioners who are familiar with basic statistical principles, including data analysis, and who wish to expand their assessment techniques to include reliability-based statistical methods.

CHAPTER 2—NOTATION

- A = observed value in Bayesian analysis
- A_s = area of tension reinforcement, in.²
- A_s' = area of compression reinforcement, in.²
- A_v = area of shear reinforcement, in.²
- a = depth of equivalent rectangular block, in.

a, b	= constants in sum of independent random variables	M	= bending moment, ft-kip
B_i	= set of all possible states of unknown parameter in Bayesian analysis, including B_j	M_A, M_{AB}, M_B	= moments at support section A, span section AB, and support section B, respectively, ft-kip
B_j	= particular state of unknown parameter in Bayesian analysis	M_D	= dead load moment (mean $\overline{M_D}$; standard deviation σ_D), ft-kip
b	= effective width of cross section, in.	M_L	= live load moment (mean $\overline{M_L}$; standard deviation σ_L), ft-kip
b_w	= web width, in.	M_n	= nominal bending moment capacity, ft-kip
C_m	= approximate moment coefficient for structural analysis	M_u	= factored applied moment, ft-kip
c	= clear concrete cover (individual measurement c_i ; mean $\bar{c}$; standard deviation s_c), in.	N	= number of events that occur in a specific time interval
c, d	= constants in product of independent random variables	n	= sample size, number of independent random variables in Taylor Series Linearization
$(\bar{c})_x$	= estimate of confidence mean cover at x confidence level, in.	P	= probability; professional factor, ratio of test to predicted value (mean value $\bar{P}$; bias coefficient δ_P ; coefficient of variation V_P)
D	= dead load effect (mean $\bar{D}$; bias coefficient δ_D ; coefficient of variation V_D)	R	= resistance (mean $\bar{R}$; standard deviation s_R ; coefficient of variation V_R)
d	= effective depth of reinforcement (mean value $\bar{d}$; coefficient of variation V_d), in.	R^2	= coefficient of determination in regression analysis
d_b	= bar diameter, in.	r_i	= rank of data point x_i
d_c	= depth of carbonation (mean $\bar{d}_c$; standard deviation s_d), in.	S	= load effect or demand (mean load effect $\bar{S}$; standard deviation σ_S ; coefficient of variation V_S)
E	= target allowable error, expressed as a fraction (or percentage) of the mean value	s	= sample standard deviation computed for a dataset; stirrup spacing, in.
e	= target allowable error	s^2	= sample variance computed for a dataset
$F(x_i)$	= sample cumulative distribution function (CDF) value for data point x_i	s_A	= standard deviation of annual maximum extreme value distribution
$F_{x_A}(x_A)$	= CDF of annual maximum extreme value distribution for data point x_A	T	= return period, years
$F_{x_N}(x_N)$	= CDF of N -year maximum extreme value distribution for data point x_N	$T_{P,n-1}$	= Student's t -distribution with $n-1$ degrees of freedom at Confidence Level P
$F_X^e(x)$	= CDF of event distribution	t_y	= time, years
$F_X^m(x)$	= CDF of extreme value distribution	V	= coefficient of variation, standard deviation divided by the mean
$f_{0.10}$	= lower 10% fractile value of a quantity	V_c	= shear resistance provided by concrete, kip
f, g	= arbitrary functions	V_{core}	= coefficient of variation of core strengths
f_c'	= concrete compressive strength (mean value $\bar{f}_c'$; coefficient of variation $V_{f_c'}$), psi	V_n	= nominal shear resistance, kip
f_{ceq}	= equivalent-to-specified concrete compressive strength, psi	V_u	= factored shear demand, kip
f_y	= reinforcement yield stress (mean value $\bar{f}_y$), psi	v	= wind speed (individual data point v_i ; mean $\bar{v}$)
f_{yeq}	= equivalent-to-specified reinforcement yield stress, psi	w	= uniformly distributed load, lb/ft; lb/ft ²
f_{yt}	= transverse reinforcement yield stress, psi	w_i	= truck weight data point (mean value $\bar{w}_e$; standard deviation s_e)
f_{yteq}	= equivalent-to-specified transverse reinforcement yield stress, psi	w_L	= uniformly distributed load due to live load, lb/ft, lb/ft ²
g_i	= standard Gumbel variate for data point x_i	w_{SDL}	= uniformly distributed load due to superimposed dead load, lb/ft; lb/ft ²
h	= overall height of rectangular or circular member, in.	w_{SW}	= uniformly distributed load due to self-weight, lb/ft; lb/ft ²
h_p	= height of beam stem projection beneath slab soffit, in.	w_u	= factored uniformly distributed load, lb/ft; lb/ft ²
h_s	= slab thickness, in.	X_i	= independent random variable (mean value $\bar{X}_i$; standard deviation σ_{X_i} ; coefficient of variation V_{X_i})
k	= carbonation rate constant		
$k_{0.10}$	= one-sided tolerance factor for 10% fractile		
k_c, k_s	= coefficient of variation modification factors specified in ACI CODE-562		
L	= live load effect (mean $\bar{L}$; bias coefficient δ_L ; coefficient of variation V_L)		
ℓ_n	= clear span length, in.		

$\bar{x}$	= sample mean value computed for a dataset
x_i	= data point in dataset
Y	= dependent random variable (mean value $\bar{Y}$; standard deviation σ_Y ; coefficient of variation V_Y)
$Z_{\alpha/2}$	= standard normal variate for Confidence Level $(1-\alpha)$
Z_i	= standard normal variate for data point x_i (predicted value $\hat{Z}_i$)
α	= parameter used to define $(1-\alpha)$ Confidence Level
α_A	= dispersion parameter for annual maximum extreme value distribution
α_N	= dispersion parameter for N -year maximum extreme value distribution
β	= reliability index
β_T	= target reliability index
Δ_M	= change of moment due to incremental change of yield stress Δf_y , ft-kip
δ	= bias coefficient, ratio of mean to nominal value
ε_i	= error in regression analysis model
ε_t	= strain in outermost reinforcing steel layer
Φ	= cumulative distribution value for standard normal distribution
Φ^{-1}	= inverse of the standard normal cumulative distribution function
ϕ	= strength reduction factor
ζ	= standard deviation of lognormally distributed population
λ	= mean value of lognormally distributed population
μ	= mean value of population
μ_A	= central tendency parameter for annual maximum extreme value distribution
μ_{max}	= upper limit of confidence interval on mean value of population
μ_{min}	= lower limit of confidence interval on mean value of population
μ_N	= central tendency parameter for N -year maximum extreme value distribution
σ	= standard deviation of population
σ_{max}	= upper limit of confidence interval on standard deviation of population
σ_{min}	= lower limit of confidence interval on standard deviation of population
σ^2	= variance of population
$\chi^2_{\alpha, n-1}$	= percentile of the χ^2 distribution with $n-1$ degrees of freedom

CHAPTER 3—ASSESSMENT

Assessment of existing structures is a multi-stage process that includes: a) review of plans and relevant documents; b) visual condition survey; c) laboratory and field testing; d) analysis and evaluation; and e) reporting (ACI PRC-364.1). This document on statistical techniques focuses on the testing portions of the assessment, how a statistically relevant program can be established, and how the testing results can be interpreted in the analysis and evaluation stage using statistical methods. Information on performing an assessment can be found in SEI/ASCE 11-99, ACI PRC-201.1, ACI PRC-364.1, ACI PRC-437, and ACI CODE-562.

3.1—Planning

It is necessary to plan the testing program before implementing it. This typically includes:

- Testing program objectives
- Parameters to be evaluated
- Testing methods to be used
- Location and number of tests
- Statistical implementation of test results

Items (a) to (c) are discussed in this chapter. Items (d) and (e) relate to the selection and implementation of the statistical techniques. Establishing the location and number of tests and the required statistical analysis of test results to achieve target confidence levels are discussed in Chapter 4.

3.2—Objective of testing program

The objective of the testing program should be determined prior to performing the testing. The objective impacts the sampling program and the type of testing.

The testing should support the overall objective of the condition assessment. The objective of the testing program should reflect the information gathered during the visual condition survey and the information determined from the available documentation.

3.3—Properties and testing

There are many parameters that may be of interest during an assessment of an existing concrete structure. These include the properties of the original concrete and reinforcement, and the properties of materials used for previous repairs or strengthening. In general, properties can be separated into three different general categories: 1) geometric; 2) strength; and 3) durability. Additional information on the applicability and limitations of testing procedures can be found in ACI PRC-214.4, ACI PRC-228.1, ACI PRC-228.2, ACI PRC-444.2, and the ASTM standard test procedures. Determining the appropriate published test procedure to evaluate a specific parameter is the responsibility of the licensed design professional performing the assessment.

3.3.1 Geometric properties—Geometric properties are frequently important for evaluation of both strength and durability and can be used to verify existing design documents when available. Typical geometric properties of interest include:

- Member dimensions (depth, width, span length)
- Concrete section loss or internal voids
- Reinforcement section loss due to corrosion
- Reinforcement quantity and placement (number of bars, bar size, slab reinforcing steel spacing, stirrup/tie spacing)
- Reinforcement cover
- Reinforcement anchorage and splice lengths
- Dimensions of previous repairs and strengthening
- Deflected shape in service
- Cracking

Some test methods that can be used to evaluate geometric properties are listed in Table 3.3.1.

3.3.2 Strength properties—Strength properties are important for performing a structural evaluation or evaluating the performance of repairs. Typical strength properties of interest include:

- (a) Concrete strength
- (b) Concrete modulus of elasticity
- (c) Mild reinforcing steel yield strength
- (d) Other reinforcement strength and properties (such as prestressing steel and fiber-reinforced polymers [FRPs])
- (e) Type and strength of existing anchorages
- (f) Bond strength of previous concrete repair applications or overlays

Table 3.3.1—Geometric test methods

Test method	Reference	Use*
Field measurements	NA	a, b, g, h, i
Exploratory openings	NA	a, c, d, e, f
Impact-echo/impulse response	ASTM C1383, ASTM C1740	a, b, g
Ground-penetrating radar	ASTM D4748, ASTM D6432	a, b, d, e, f, g
Covermeter	NA	d, e
Infrared thermography	ASTM D4788	b
Sounding	ASTM D4580	b, g
Reinforcing steel measurements	ACI PRC-364.14	c

*Refers to list of geometric properties in 3.3.1; NA is not available.

Table 3.3.2—Strength test methods

Test method	Reference	Use*
Concrete compressive strength	ASTM C39/C39M, ASTM C42/C42M	a
Splitting tensile strength	ASTM C496/C496M	a
Concrete modulus of rupture	ASTM C78/C78M	a
Static modulus of elasticity	ASTM C469/C469M	b
Bond pulloff testing	ASTM C1583/C1583M	f
Mechanical testing of steel	ASTM A370	c, d
FRP tensile strength	ASTM D7565/D7565M	d
Load testing	ACI CODE-437.2	d, e, g

*Refers to list of strength properties in 3.3.2.

Table 3.3.3—Durability test methods

Test method	Reference	Use*
Air-void system	ASTM C457/C457M	a
Petrographic examination	ASTM C856/C856M	a, c, d
Aggregate petrographic examination	ASTM C295/C295M	d
Acid-soluble chloride content	ASTM C1152/C1152M	b
Water-soluble chloride content	ASTM C1218/C1218M	b
Phenolphthalein spray	NA	c
Half-cell potential	ASTM C876	e
Corrosion rate	NA	e
Rebound (Schmidt) hammer	ASTM C805/C805M	f

*Refers to list of durability properties in 3.3.3; NA is not available.

(g) Strength of the structure or specific elements

Some test methods that can be used to evaluate strength properties are listed in Table 3.3.2.

3.3.3 Durability properties—Durability properties are frequently important as part of the assessment and can be used to evaluate the type and extent of repairs required, as well as the potential for future deterioration. Typical durability properties of interest include:

- (a) Resistance to freezing and thawing
- (b) Concrete chloride profiles
- (c) Concrete carbonation depth
- (d) Reactive aggregate presence
- (e) Corrosion state
- (f) Extent of fire damage

Some test methods that can be used to evaluate durability properties are listed in Table 3.3.3.

CHAPTER 4—BACKGROUND ON STATISTICAL TECHNIQUES

The primary objective of this chapter is to provide background on statistical techniques used in assessment of existing concrete structures. Detailed examples of the application of these techniques are provided in **Chapters 5 and 6**. More detail on these techniques, which apply only to unbiased samples obtained randomly, is presented in **Appendix A**.

It is beneficial to define a sampling framework to organize the data gathering and statistical assessment process. The sampling framework should:

- (a) Define the purpose of the sampling (for example, what question(s) is the assessment trying to solve? Refer to **Section 3.2**)
- (b) Identify critical parameters to be sampled (**Section 3.3**)
- (c) Define the population, potential sources of sampling bias, and the numbers of samples to be obtained (**Section 4.1**)
- (d) Establish how the collected data will be evaluated (**Sections 4.2 through 4.5**).

4.1—Defining the population and determining sample sizes

This section presents a basis for defining the population to be investigated and illustrates various options for determining the sample size.

Defining the population and sample size to yield the desired accuracy of sampled parameters is an essential initial step that requires engineering judgment. Defining the population, discussed in 4.1.1, is an important aspect, but selecting the required sensitivity is nearly as important. Increasing the number of samples narrows the uncertainty of parameters and increases confidence, but also inevitably increases the cost of the assessment, although it may reduce the cost of the final solution. The necessary accuracy of each parameter should depend on its relative importance in the assessment. For example, the flexural capacity of a reinforced concrete beam is relatively insensitive to the concrete compressive strength. It is therefore appropriate to determine the concrete compressive strength approximately with relatively few samples. In contrast, the flexural strength of the

beam is sensitive to the yield strength and effective depth of the reinforcement, so it is desirable to determine these quantities more accurately using larger sample sizes. Extracting reinforcing bar samples for testing is typically expensive, so it may be prudent to focus on accurately determining the effective reinforcement depth, particularly in a thin slab.

4.1.1 Defining the population—A key component is to identify the “population” used to quantify statistical parameters, such as mean values or standard deviations. The objective is to limit the population by defining the problem and the essential data needed to solve it. There are no fixed or defined rules for doing this, but suitable procedures can be developed on a case-by-case basis given the prior information the licensed design professional knows about the structure and the nature of the problem to be solved.

The sample size and the associated investigation effort is highly dependent on the size of the population because the number of samples needed to achieve a certain confidence, or accuracy, increases as the population size increases. The population size may be infinite—consider, for example, a population representing the clear cover to the ties in columns of a large, multi-story building. To this end, limiting the population by defining the problem allows a relatively small sample to be both representative and sufficient.

Consideration of the prior information includes the following:

(a) **Overall construction history:** Was the structure constructed as a single initiative by a single contractor, or have additions been added by different contractors over the years? In the former case, the entire structure could be classified as one population because the material properties and dimensional tolerances are likely to be consistent throughout the structure. In the latter case, the original structure and each addition should be considered as distinct populations with, for example, unique material properties, unique construction tolerances, unique deficiencies, or unique deterioration. Subdividing a population into distinct subpopulations inevitably increases the total number of specimens needed to evaluate a property, as the property must be evaluated for each subpopulation.

(b) **Available information:** Do the original construction documents, if available, specify different material grades for different members (or for longitudinal and transverse reinforcement in members)? If so, each region of the structure with a unique material grade should be classified as a distinct population.

Consideration of the nature of the problem to be solved includes the following:

(a) Is the purpose of the assessment to determine whether a specific defect affects the structural reliability or durability, or to determine the structural reliability or durability in the context of a proposed change of occupancy or environmental exposure? In the former case, the population is limited to the region(s) where the defect(s) is present. In the latter case, it may be appropriate to classify the entire structure as a single population. If the change in occupancy or environmental exposure pertains to one region of the structure only, then that region should be considered a distinct population.

(b) Does the preliminary investigation (for example, visual inspection, preliminary assessment calculations) reveal problematic conditions in specific areas? This would allow the definition of two distinct populations: one representing “problematic” conditions or locations and the other representing “non-problematic” conditions or locations.

(c) Do the environmental exposure conditions or extents of deterioration vary within the structure or between structures? If so, it may be appropriate to split the elements or structures into different populations for evaluation. If not, it may be appropriate to classify the entire structure or group of structures as a single population.

The process of defining the populations to be sampled provides an opportunity to consider possible sources of sampling bias. Bias is introduced when the selection of samples is not representative of the population being sampled. For example, sampling only in regions where deterioration is visibly present and ignoring undeteriorated regions creates bias. Bias associated with location can be minimized by characterizing the volume, or surface area, using equally sized elements, numbering the elements sequentially, and then using a random number generator to select the elements to be sampled.

Bias can be eliminated if the licensed design professional can use knowledge to define different populations of the structure based on observed differences in damage and/or performance. As an example, the underwater part of a coastal foundation is found to have lost substantial cover and shows aggregate while the above-water section shows no such deterioration. The licensed design professional suspects that the underwater section has been subjected to sulfate attack because of its submergence in seawater, and they may reasonably test only the deteriorated material for sulfate attack. The underwater portion is a different population than the above-water portion.

4.1.2 Determining sample size—Classical and specialized statistical methods are available to determine sample sizes. Practically, however, these methods require modification, using engineering judgment, to suit project-specific constraints and conditions.

The classical statistical approaches are presented in statistics textbooks (Ang and Tang 1975; Larsen and Marx 2017) and ASTM E122. If the population is large, normally distributed, and its variability is known, then the required number of samples or test specimens, n , can be calculated as

$$n = \left(z_{\alpha/2} \cdot \frac{\sigma}{e} \right)^2 \quad (4.1.2a)$$

where $z_{\alpha/2}$ is the inverse of the standard normal distribution; $(1-\alpha)$ is the desired confidence level; σ is the standard deviation of the population; and e is the target allowable error. The value of $z_{\alpha/2}$ can be obtained from tables in introductory statistics textbooks (Larsen and Marx 2017) or using the Excel spreadsheet function “NORM.S.INV(P),” where P , the target probability, is equal to $\alpha/2$. The selection of a suitable confidence level is discussed in 4.4.

For example, assume a population of clear concrete covers has a mean of 2.0 in. and a standard deviation of 0.4 in. The number of samples required to achieve an estimate of the mean within ± 0.1 in. at the 90% Confidence Level ($\alpha = 0.10$, $z_{\alpha/2} = z_{0.05} = \text{NORM.S. INV}(0.05) = 1.644$) is

$$n = \left(z_{\alpha/2} \cdot \frac{\sigma}{e} \right)^2 = \left(1.644 \cdot \frac{0.4 \text{ in.}}{0.1 \text{ in.}} \right)^2 = 43$$

Equation (4.1.2a) can be rewritten in terms of the coefficient of variation, V , which is the dimensionless ratio of the standard deviation divided by the mean, as

$$n = \left(z_{\alpha/2} \cdot \frac{V}{E} \right)^2 \quad (4.1.2b)$$

where E is the error expressed as a fraction of the mean. Errors attributable to the accuracy of the testing method also contribute to E . The target error is defined by the licensed design professional and can be taken as the tolerance specified in **ACI SPEC-117**. For the values from the previous example, $V = (0.4 \text{ in.}/2 \text{ in.}) = 0.20$, $E = (0.1 \text{ in.}/2 \text{ in.}) = 0.05$; the number of samples is

$$n = \left(z_{\alpha/2} \cdot \frac{V}{E} \right)^2 = \left(1.644 \cdot \frac{0.20}{0.05} \right)^2 = 43$$

When the standard deviation or coefficient of variation of the population is unknown, it is necessary to assume a value to use the classical approach that requires subsequent verification by sampling. This is often impractical.

Specialized statistical approaches that combine random sampling with probabilistic approaches—for example, involving the hypergeometric distribution—can be used to reduce the number of required samples. Such methods are beyond the scope of this report but are well-described elsewhere (**Ang and Tang 1975**; **Theryo et al. 2013**).

It is often necessary in practice to reduce the sample size to suit project-specific considerations and constraints using engineering judgment. This approach may be necessary if the population size is limited, as discussed in **Chapter 5**, or the cost to perform the assessment is prohibitive. The sample mean and sample standard deviation can be bounded to achieve desired confidence levels, as illustrated in Section 4.4, particularly for small sample sizes.

4.1.3 Example calculation—An assessment team plans to measure the average clear concrete cover to the stirrups of beams in a building. In this case, the population consists of all 400 beams in the seven-story building. The population mean and standard deviation are unknown, but the design cover is believed to be 2 in. Use Eq. (4.1.2b) to determine the necessary sample size to be 90% confident of achieving a target error E of 5 or 10% of the mean value if the coefficient of variation of the underlying population varies between 10 and 40%.

Table 4.1.3a—Required number of samples at 90% Confidence Level for given target error E and coefficient of variation V

Target error E	Assumed coefficient of variation V				
	10%	15%	20%	30%	40%
$\pm 5\%$	11	24	43	97	173
$\pm 10\%$	3	6	11	24	43

Table 4.1.3b—Required number of samples at 75% Confidence Level for given target error E and coefficient of variation V

Target error E	Assumed coefficient of variation V				
	10%	15%	20%	30%	40%
$\pm 5\%$	5	12	21	48	85
$\pm 10\%$	1	3	5	12	21

This calculation has been partially illustrated in the previous example calculation where, for a 90% Confidence Level, it was shown that 43 specimens are necessary if the population V is 20% and the target E is 5%. Other sample sizes for various populations V and target errors E are shown in Table 4.1.3a. Clearly, the required number of samples determined using Eq. (4.1.2a) and (4.1.2b) can become very large if the assumed V is high and the allowed error is low; for example, $n = 173$ for $V = 40\%$ and $E = 5\%$.

If the confidence level is reduced to 75%, the required number of samples is reduced, as shown in Table 4.1.3b. For the 75% Confidence Level, $\alpha = 0.25$, $z_{\alpha/2} = z_{0.125} = 1.150$. For this case, the necessary sample size at the 75% Confidence Level is approximately half as large as that for the 90% Confidence Level.

At the start of an assessment, the licensed design professional may conservatively select a large coefficient of variation to determine the required sample size. It is sometimes beneficial to reevaluate the required sample size during the assessment and update the V -values based on field measurements.

4.2—Characterizing data

This section describes how to characterize data using the sample statistics, specifically the sample mean and sample standard deviation, and how to determine the probable distribution type using probability paper.

Figure 4.2 shows the probability density functions (PDFs) for normal, logarithmic normal (or lognormal) and extreme value Type I distributions. In all cases, the mean value is 100 and the standard deviation is 25, so the coefficient of variation is 0.25. The normal distribution graphs as the familiar, symmetrical bell curve. Dimensions, concrete strengths, and quantities that can be expressed as the sum of independent random variables are typically normally distributed. The lognormal distribution is not symmetric, with a lighter lower tail and a heavier upper tail than the normal distribution. Steel yield strengths, quantities that are always positive, and quantities that can be expressed as the product of independent random variables typically have

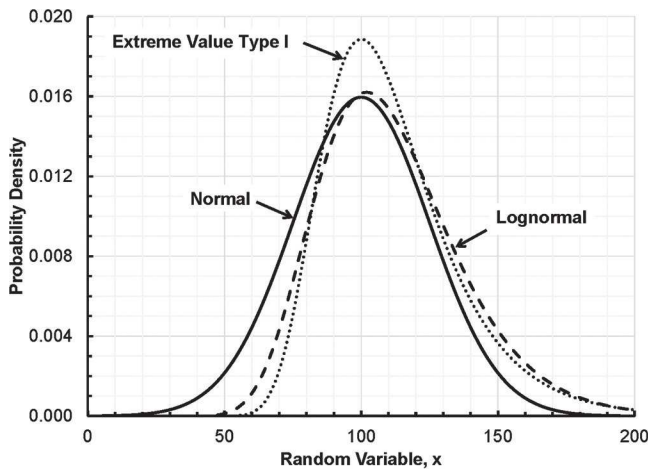


Fig. 4.2—Probability density functions for normal, lognormal, and extreme value Type I distributions.

lognormal distributions. The extreme value Type I distribution is also unsymmetric, with the lightest lower tail and an upper tail that is similar to that of the lognormal distribution. As described in A.6, extreme value Type I distributions are often used to characterize the maximum transient loads that may be expected during the lifetime of the structure. Given the similarity of the upper tails of the lognormal and extreme value Type I distributions, it is also possible to characterize maximum transient loads using lognormal distributions.

4.2.1 Computing sample statistics—Most statistical analyses start with the computation of the sample mean value $\bar{x}$ and sample standard deviation s of the sample data. To compute the mean value, use the “AVERAGE(●)” function in an Excel spreadsheet, where ● is the range of cells that contain the data, or

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (4.2.1a)$$

where x_i are the individual sample results; and n is the sample size.

To compute the sample standard deviation, use the “STDEV.S(●)” function in an Excel spreadsheet or

$$s = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (4.2.1b)$$

4.2.2 Determining type of distribution using probability paper—The distribution type is often readily determined from the sample cumulative distribution of the data. The data are first ordered from smallest to largest and assigned rank numbers r_i , from 1 for the smallest measured value to n for the largest. The sample cumulative distribution function (CDF) value $F_X(x_i)$, which is the probability that the random variable has a value less than or equal to x_i , for each data point is computed as

$$F_X(x_i) = \frac{r_i}{n+1} \quad (4.2.2a)$$

To determine if the data are normally distributed, these sample CDF values are transformed to equivalent standardized normal variates Z_i —that is

$$Z_i = \Phi^{-1}[F_X(x_i)] \quad (4.2.2b)$$

where Φ^{-1} is the inverse of the standard normal cumulative distribution function. In an Excel spreadsheet, this function is “NORM.S.INV(●)”, where (●) is the cell that contains $F_X(x_i)$. Then, a graph is created on “probability paper,” with x_i as the independent variable and Z_i as the dependent variable. If the relationship is reasonably linear, it can be assumed that the data are normally distributed. An example in 4.2.4 illustrates this procedure.

To determine if the data are lognormally distributed, the same steps are taken but each Z_i is graphed against $\ln(x_i)$. If the relationship is reasonably linear, it can be assumed that the data are from a lognormal distribution.

To determine if the data are from an extreme value Type I distribution, the sample CDF values are computed using the aforementioned procedure and Eq. (4.2.2a). The standard Gumbel variate g_i is computed as

$$g_i = -\ln\{-\ln[F_X(x_i)]\} \quad (4.2.2c)$$

Then, a graph is created with x_i as the independent variable and g_i as the dependent variable. If the relationship is reasonably linear, it can be assumed that the data follow an extreme value Type I distribution. There is no Gumbel variate calculation in Excel, so Eq. (4.2.2c) must be used. Example A.10.3 gives an illustration of this procedure.

4.2.3 Identification of outliers—An outlier is a data point that appears to deviate significantly from other data in the sample. Ideally, the decision to remove a suspected outlier from a dataset is based on some physical explanation for the deviation. For example, a strength test result may be unrepresentative if there is clear evidence that the specimen was damaged before being tested, improperly mounted in the testing machine, or tested at an inappropriate rate of loading.

Statistical methods for identification of outliers are presented in ASTM E178. They are not particularly effective for small sample sizes (Bartlett and MacGregor 1995).

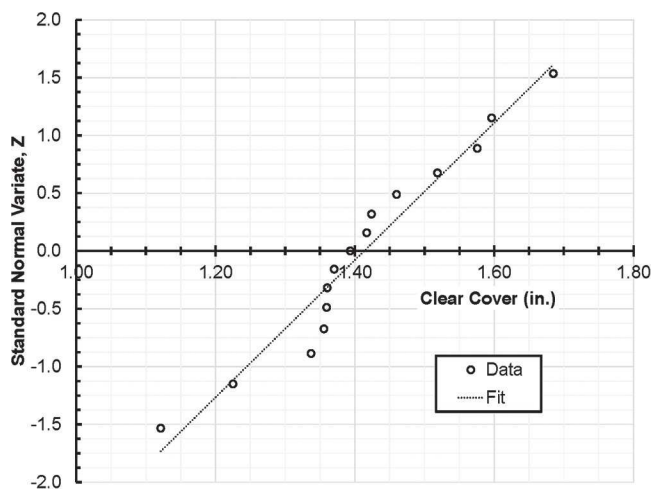
4.2.4 Calculation example—The left column in Table 4.2.4 presents 15 measurements of the clear concrete cover for columns in a building. Determine the sample mean and sample standard deviation for these data and graph the sample cumulative distribution on normal probability paper.

Using the Excel functions “AVERAGE(●)” and “STDEV.S(●)”, the sample mean of the 15 values is computed as 1.413 in. and the sample standard deviation as 0.143 in.

The cover values will be the independent variables on the probability paper plot. As shown in Table 4.2.4, the data are sorted from least (1.12 in.) to largest (1.68 in.) and assigned a rank from 1 for the least value, to 15. The same CDF value $F_X(x_i)$ is computed using Eq. (4.2.2a); for example, for $r_i = 4$, $F_X(x_i) = 4/(1 + 15) = 0.25$. These sample CDF values are

Table 4.2.4—Analysis of in-place cover measurements

Cover c_i , in.	Rank	$F_X(x_i)$	Z_i	Fitted
1.12	1	0.0625	-1.534	-1.733
1.23	2	0.125	-1.150	-1.116
1.34	3	0.1875	-0.887	-0.452
1.36	4	0.25	-0.674	-0.342
1.36	5	0.3125	-0.489	-0.319
1.36	6	0.375	-0.319	-0.314
1.37	7	0.4375	-0.157	-0.255
1.39	8	0.5	0.000	-0.117
1.42	9	0.5625	0.157	0.022
1.42	10	0.625	0.319	0.063
1.46	11	0.6875	0.489	0.277
1.52	12	0.75	0.674	0.625
1.58	13	0.8125	0.887	0.965
1.60	14	0.875	1.150	1.086
1.68	15	0.9375	1.534	1.612

**Fig. 4.2.4—Concrete cover data graphed on normal probability paper.**

computed to standard normal variates Z_i using the Excel “NORM.S.INV(●)” function.

The best fit to these data is obtained by regression analysis, with Z_i as the dependent (or Y-axis) variable, and the cover measurements c_i as the independent (or X-axis) variable. The regression analysis yields

$$\hat{Z}_i = -8.385 + 5.933c_i$$

where $\hat{Z}_i$ is the predicted value of Z_i for a given c_i . As c_i is measured in inches, the coefficient 5.933 obtained by regression must have the unit 1/in. The mean cover corresponds to $\hat{Z}_i = 0$; therefore, $-(-8.385/5.933) = 1.413$ in. The standard deviation is the horizontal distance corresponding to a change in $\hat{Z}_i$ of 1.0; therefore, $(1/5.933) = 0.169$ in. The sample mean is identical to that computed using the “AVERAGE(●)” Excel function and the sample standard

deviation is slightly greater than that computed using the “STDEV.S(●)” Excel function.

Figure 4.2.4 shows the data and fitted CDF on normal probability paper. The data define a reasonably linear relationship—increasing the sample size would clarify this—and so the cover can be assumed to be normally distributed. The extreme high and low values lie close to the fitted line and thus are unlikely to be outliers.

4.3—Comparing and combining datasets

It is advantageous to combine data from different sources if the statistical parameters of the data from each source are similar. The resulting dataset is larger, so inferences made from it are more accurate and more robust. This section describes how the sample mean values and sample standard deviations of different datasets can be compared.

4.3.1 Comparing sample mean values of different datasets—The sample means of two datasets can be compared using the Student’s t -test to determine if their difference is statistically significant, and not simply a value that might occur by chance given the scatter of the data. The t -test is derived assuming the underlying populations are normally distributed. It is desirable, though not necessary, that the sample sizes of the two datasets are similar.

Details of the t -test are presented in any introductory statistical textbook (Larsen and Marx 2017). The test is also available as the Excel spreadsheet function “T.TEST(array1, array2, tails, type)”, where array1 and array2 are the two datasets, respectively; tails distinguishes between a one-tailed test and a two-tailed test; and type is the type of test carried out. A two-tailed test is typically conducted to check whether the sample mean of the first dataset is significantly greater than or significantly less than the sample mean of the second dataset. The one-tailed test, used less frequently, requires the analyst to assume that the first mean must be greater than (or must be less than) the second mean and tests only if this difference is statistically significant. The type distinguishes between cases where the standard deviations of the two datasets are similar or differ significantly, which can be determined following the procedure described in 4.3.2.

The “T.TEST(●)” function returns a P -value, which is the probability of the difference of the sample mean values occurring naturally, or by chance, given the variability of the datasets. Conventionally, a difference is deemed to be statistically significant if its P -value is less than 0.05, or 5%.

4.3.2 Comparing sample standard deviations of different datasets—The sample standard deviations of two datasets can be compared using the f -test to determine if the difference in the sample variances is statistically significant. The sample variance is simply the square of the sample standard deviation and thus is different from the dimensionless coefficient of variation. The f -test is also derived assuming the underlying populations are normally distributed. It is again desirable, but not necessary, that the sample sizes of the two datasets are similar.

Details of the f -test are again presented in any introductory statistical textbook. The test is also available as an Excel

spreadsheet function “F.TEST(array1, array2), where array1 and array2 are the two datasets, respectively.

The “F.TEST(●)” function returns a P -value, which is the probability of the difference of the sample variances occurring naturally given the variability of the datasets. Again, the difference is conventionally assumed to be statistically significant if the P -value is less than 0.05 or 5%.

Usually the f -test is conducted before the t -test in data analysis. If the f -test indicates that the difference in sample variances is statistically significant, the t -test is performed with type = 3 to account for this significant difference. If the f -test indicates that the difference in sample variances is not statistically significant, the t -test is performed with type = 2 to reflect this.

4.3.3 Example calculation—Table 4.3.3 shows the concrete compressive strengths obtained from 12 cores drilled from walls and 12 cores drilled from slabs of an existing concrete building. The sample means and sample standard deviations are computed as shown. Is the observed 405 psi difference between the average wall core strength and the average slab core strength statistically significant?

This problem is readily solved using the Excel spreadsheet “T.TEST(●)” and “F.TEST(●)” functions. In both cases, “array1” is simply the set of 12 wall core test strengths and

“array2” is the set of 12 slab core test strengths. It is not necessary that the number of data points in “array1” and “array2” are equal.

The f -test is performed first. The ratio of the sample variances seems large: the ratio of the observed wall strength variance to the observed slab strength variance is $(317/186)^2 = 2.9$. The F.TEST(●) function returns a P -value of 0.0919, that exceeds 0.05 and so indicates that the difference in the sample variances (or sample standard deviations), although large, is not statistically significant.

A two-tailed t -test is therefore performed with the “type” variable set equal to 2: the Excel function is “T.TEST(wall dataset, slab dataset, 2, 2)”. It returns a P -value of 0.00095, which is less than 0.05 and thus indicates that the observed 405 psi difference between the sample mean values is statistically significant. It would therefore not be appropriate to combine the wall core strength data with the slab core strength data to create a single dataset.

If the two-tailed t -test is performed with the “type” variable set equal to 3, which is appropriate if the sample variances are not equal, it returns a P -value of 0.0013. This value is greater than that obtained previously, but still markedly less than 0.05 and so confirms that the difference in sample mean cores strengths is statistically significant. This is therefore a true difference, not a difference that would be expected to occur naturally by chance, so the datasets should not be combined.

Figure 4.3.3 shows the slab and wall data on a single figure to facilitate a visual comparison of the two datasets. Clearly, all but the highest slab core strengths are less than all but the lowest wall core strengths. This conclusion is consistent with the observation that slabs and walls are usually placed separately, often have different mixture proportions, and are cured differently, so are typically distinct populations.

4.4—Bounding sample statistics using confidence intervals

This section describes how to derive upper and lower bounds on the population mean and population standard deviation computed from a set of data using confidence intervals. Because the data represent a discrete, often small sample of the population, the sample mean $\bar{x}$ and sample standard deviation s typically differ from the population mean μ and standard deviation σ . As the sample size increases, the sample statistics tend to approach the population values.

The bounds on the population mean or population standard deviation are defined using confidence intervals to achieve a desired level of accuracy. Two-sided confidence intervals define the upper and lower bounds of a parameter. As an example, choosing a 90% Confidence Level, upper and lower bounds can be defined such that there is 90% confidence that the interval between the bounds contains the population mean or population standard deviation.

Often, the upper or lower limit of a quantity is desired—for example, the upper limit of a demand or the lower limit of a capacity. In this case, a one-sided confidence interval is used. For the 90% Confidence Level, there is a 90% chance

Table 4.3.3—Wall and slab core strength data

Core number	Wall core strength, psi	Slab core strength, psi
1	3710	2680
2	3630	3080
3	2870	3050
4	3180	3040
5	3590	3170
6	3330	3050
7	4160	2960
8	3560	2810
9	3370	3150
10	3280	3430
11	3450	3150
12	3340	3040
Average	3456	3051
Standard deviation	317	186

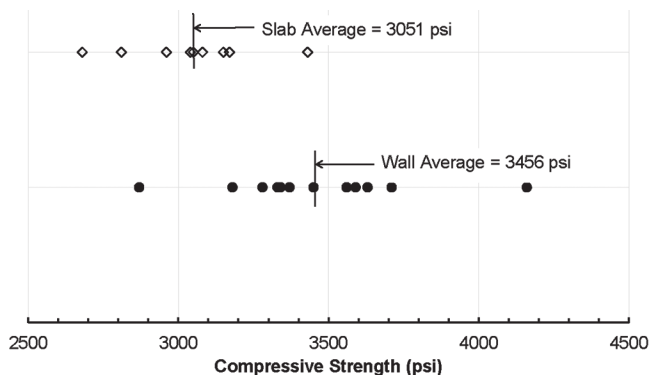


Fig. 4.3.3—Wall and slab core strength data.

that the true parameter is less than the upper limit or a 90% chance that the true parameter is greater than the lower limit.

This section presents a discussion on confidence level selection and methods for computing bounds on the sample mean and the sample standard deviation for a given sample size to achieve a desired confidence level. The application of these approaches is illustrated with a calculation example in 4.4.4.

4.4.1 Confidence level selection—Confidence level selection should consider at least the following factors:

(a) **Severity, or consequence, of failure:** The selected confidence level should reflect the risk associated with continued deterioration or failure. For example, if the failure of an existing structure would result in a significant threat to public safety, there would be higher consequences of failure than a structure whose failure would have a minor impact on public safety. As such, a higher confidence level, and subsequently a higher frequency of testing, should be selected for the first structure. When evaluating the consequence of failure, all hazard sources should be evaluated, including threat to public safety, equipment damage, environmental damage, reduced function, reduced serviceability, increased future maintenance or repair costs, and negative impact to reputation.

(b) **Likelihood of failure:** The selected confidence level should consider the current condition of the structure and the likelihood of its failure under service loads. For example, if an existing structure exhibits widespread spalling and exposed reinforcing steel, there could be a higher likelihood of failure within the target service life of the repair than for a structure with minor localized cracking. A higher confidence level should be considered for the more damaged structure to identify the root causes of deterioration and properly address them in the repair design.

(c) **The sensitivity or the importance of the parameter being measured:** For example, consider a large unreinforced hydroelectric dam with a high consequence of failure and no reinforcement. Given that the durability of the dam is not sensitive to carbonation depth, a high confidence level with regard to carbonation depth may not be required. Conversely, consider a heavily reinforced slab on ground with shallow cover and high chloride concentration. Although the consequences of failure are low, a high confidence level may be desired given the sensitivity of the slab durability to carbonation depth.

It is important for the licensed design professional and owner to determine appropriate confidence levels collaboratively. Engineering input should include evaluation of likelihood of failure, technical considerations regarding typical degradation rates for various types of damage mechanisms, variability of measurement techniques, redundancy, and strength and durability implications. The licensed design professional should obtain feedback from the owner concerning the intended use of the structure, required service life, cost of assessment, and tolerance for the consequences of failure. ACI PRC-228.1 recommends confidence levels of 95%, 90%, and 75% for estimating concrete strength in existing structures. However, after consultation with the

owner, the licensed design professional may decide to use other confidence levels, if appropriate, depending on the measured parameter and project specifics.

4.4.2 Confidence interval on the population mean—The upper and lower bounds on the mean value of a normally distributed population, μ_{max} and μ_{min} , respectively, are given by

$$\mu_{max} = \bar{x} + \frac{T_{\alpha/2, n-1} s}{\sqrt{n}} \quad (4.4.2a)$$

and

$$\mu_{min} = \bar{x} - \frac{T_{\alpha/2, n-1} s}{\sqrt{n}} \quad (4.4.2b)$$

where $\bar{x}$ is the sample mean; s is the sample standard deviation; n is the sample size; and $T_{\alpha/2, n-1}$ is the $\alpha/2$ percentile of the student's t -distribution. The confidence level is simply $(1 - \alpha)$, so, for example, for a 90% Confidence Level, $\alpha = 10\%$ or 0.10, and $\alpha/2 = 0.05$. The $T_{\alpha/2, n-1}$ value can be obtained from tables in any elementary statistics textbook (Larsen and Marx 2017) or using the Excel spreadsheet function “T.INV.2T(α , $n - 1$)”.

Similarly, the one-sided upper and lower bounds on the mean value are given by

$$\mu_{max} = \bar{x} + \frac{T_{1-\alpha, n-1} s}{\sqrt{n}} \quad (4.4.2c)$$

and

$$\mu_{min} = \bar{x} - \frac{T_{1-\alpha, n-1} s}{\sqrt{n}} \quad (4.4.2d)$$

where $T_{1-\alpha, n-1}$ is the α percentile of the Student's t -distribution determined using the Excel spreadsheet function “T.INV($1 - \alpha$, $n - 1$)”.

Equations (4.4.2a) through (4.4.2d) are exact when the underlying population is normal, and approximate, but reasonably informative otherwise.

4.4.3 Confidence interval on the population standard deviation—The two-sided confidence interval and one-sided confidence limit on the standard deviation can be determined using a similar approach involving the χ^2 distribution, which pertains to the variance, instead of the t -distribution, which pertains to the mean. The upper and lower bounds on the standard deviation of a normally distributed population, σ_{max} and σ_{min} , respectively, are given by

$$\sigma_{max} = \sqrt{\frac{(n-1)s^2}{\chi_{\alpha/2, n-1}^2}} \quad (4.4.3a)$$

and

$$\sigma_{min} = \sqrt{\frac{(n-1)s^2}{\chi_{1-\alpha/2, n-1}^2}} \quad (4.4.3b)$$

where $\chi_{\alpha/2, n-1}^2$ and $\chi_{1-\alpha/2, n-1}^2$ are the $\alpha/2$ and $(1 - \alpha/2)$ percentiles of the χ^2 distribution. These values are tabulated in

introductory statistical textbooks or can be obtained using the Excel spreadsheet functions “CHISQ.INV($\alpha/2$, $n - 1$)” and “CHISQ.INV($1 - \alpha/2$, $n - 1$)”, respectively.

The one-sided upper-bound confidence limit on the standard deviation may be useful because it represents an upper bound on the uncertainty of a quantity. It is computed as

$$\sigma_{max} = \sqrt{\frac{(n-1)s^2}{\chi_{\alpha,n-1}^2}} \quad (4.4.3c)$$

Again, the equations are exact when the underlying population is normal, and approximate, but reasonably informative otherwise.

4.4.4 Calculation example—For the 12 wall core strengths in Table 4.3.4, determine:

- (a) The 95% Confidence Interval on the mean
- (b) The 95% lower-bound Confidence Limit on the mean
- (c) The 90% upper-bound Confidence Limit on the standard deviation

(a) The wall core strengths shown in Table 3.3.4 have $\bar{x} = 3456$ psi, $s = 317$ psi, and $n = 12$. For a 95% Confidence Interval, $\alpha = 0.05$ and $T_{\alpha/2,n-1} = T_{(0.025),11} = T.INV.2T(0.05,11) = 2.201$, so, from Eq. (4.4.2a) and (4.4.2b)

$$\mu_{max} = \bar{x} + \frac{T_{\alpha/2,n-1}s}{\sqrt{n}} = 3456 + \frac{2.201(317)}{\sqrt{12}} = 3657 \text{ psi}$$

$$\mu_{min} = \bar{x} - \frac{T_{\alpha/2,n-1}s}{\sqrt{n}} = 3456 - \frac{2.201(317)}{\sqrt{12}} = 3254 \text{ psi}$$

Thus, there is 95% confidence that the population mean lies between approximately 3250 and 3660 psi.

(b) The 95% lower-bound Confidence Limit on the mean, given $T_{1-\alpha,n-1} = T_{(1-0.05),11} = 1.796$, is from Eq. (4.4.2d)

$$\mu_{min} = \bar{x} - \frac{T_{1-\alpha,n-1}s}{\sqrt{n}} = 3456 - \frac{1.796(317)}{\sqrt{12}} = 3291 \text{ psi}$$

Note also that the Excel function T.INV(0.05,11) returns a negative quantity, -1.796 , so the sign for the second term when computing μ_{min} in Eq. (4.4.2d) has to change

$$\mu_{min} = \bar{x} + \frac{T.INV(0.05,11)s}{\sqrt{n}} = 3456 + \frac{-1.796(317)}{\sqrt{12}} = 3291 \text{ psi}$$

Thus, the 95% lower-bound Confidence Limit on the mean is approximately 3290 psi. Note that because this is a single limit, not a range, the value is slightly greater than that obtained in (a).

The durability example in Chapter 5 features calculations of both upper-bound and lower-bound confidence limits.

(c) The 90% upper-bound Confidence Limit on the standard deviation, given $\chi_{\alpha,n-1}^2 = \chi_{0.10,11}^2 = \text{CHISQ.INV}(0.10,11) = 5.578$, is from Eq. (4.4.3c)

$$\sigma_{max} = \sqrt{\frac{(n-1)s^2}{\chi_{\alpha,n-1}^2}} = \sqrt{\frac{(11)317^2}{5.578}} = 445 \text{ psi}$$

Thus, the 90% upper Confidence Limit on the standard deviation is approximately 450 psi.

4.4.5 Extension of calculation example from 4.1.3—Despite completing the sample size calculations summarized in Table 4.1.3, access constraints and budget limitations caused the licensed design professional to measure the clear concrete cover to the stirrups in 21 beams. The full length of each beam was scanned, and a total of 253 stirrups were detected in the 21 beams. The data are reported as the average minimum clear cover measurements for each beam and as individual minimum clear cover measurements for each stirrup. Compute the 90% Confidence Intervals on the mean clear concrete cover and the cover standard deviation defining the population to be represented by the beam measurements or by the stirrup measurements.

Table 4.4.5 presents the various relevant quantities. The sample mean values for the beams and individual stirrups are similar. The sample standard deviations are markedly greater for the individual stirrups than for the beams, and the spread between maximum and minimum values is greater for the stirrups. However, the 90% Confidence Intervals on the mean value and particularly the standard deviation are tighter for the individual stirrups than for the beams because the number of measurements is greater. These confidence intervals are computed as illustrated in 4.4.4. The sample coefficient of variation is 17% for the beam population and 22% for the stirrup population. Due to the greater sample size, however, the population mean can be estimated to lie within a 6% interval for the 21 beam samples and within a 2% interval for the 253 stirrup samples.

This comparison assumes the stirrups are randomly sampled (as a population). In fact, they were not—they

Table 4.4.5—Population statistics for different population definition

Parameter	Population	
	Beams	Stirrups
Number of samples n	21	253
Total population	400	Effectively infinite
Maximum cover $c_{i,max}$, in.	1.84	2.29
Minimum cover $c_{i,min}$, in.	1.04	0.76
Sample mean $\bar{c}$, in.	1.33	1.32
Sample standard deviation s_c , in.	0.22	0.29
90% Confidence interval on mean, in.	1.25 to 1.41	1.29 to 1.35
90% Confidence interval on standard deviation, in.	0.18 to 0.30	0.27 to 0.31

were sampled in groups. However, treating the samples as normally distributed is reasonable in this context.

The results show that no one beam is much different than another; therefore, the results from the larger dataset—that is, the 253 stirrup measurements—should be used to determine the mean and standard deviation for the evaluation. In general, use of a larger dataset is preferred if individual groups are not significantly different.

4.5—Determining fractile values

This section briefly defines fractile values, or “fractiles”, and outlines methods for computing fractiles from datasets.

A fractile of a quantity is determined using the CDF of that quantity. For example, assume the CDF value corresponding to a yield stress of 58 ksi for Grade 60 reinforcing steel equals 0.10, indicating that the probability that the actual yield stress is less than 58 ksi is 0.10, or 10%. The 10% yield stress fractile for Grade 60 reinforcing steel is therefore 58 ksi.

Lower fractiles, often the 5% or 10% fractile, are typically used to quantify material strength parameters as well as resistances. Upper fractiles can be used to quantify loads, although they are more typically characterized using extreme value distributions as described in 4.6. Upper fractiles are useful to quantify other parameters where upper bounds are desirable, such as stirrup spacing or carbonation depth.

The fractile value is computed as some number of standard deviations above or below the mean value. For example, for a normally distributed quantity, the lower 10% fractile, $f_{0.10}$, is 1.282 standard deviation σ , below the mean value μ

$$f_{0.10} = \mu - 1.282\sigma \quad (4.5)$$

Fractile values determined from data reflect the uncertainty of the sample mean $\bar{x}$ and the uncertainty of the sample standard deviation s , and thus require larger confidence intervals than either the mean, as described in 4.4.2, or the standard deviation, as described in 4.4.3.

The strength assessment example in [Chapter 6](#) includes the approximate calculation of lower fractiles of the concrete compressive strength and the reinforcing steel yield strength. Refer also to [ACI PRC-214.4](#).

4.6—Characterizing data representing extreme transient loads

This section briefly summarizes the characterization of data representing transient loads. Elaboration on this topic, including mathematical procedures to manipulate transient load data and statistical parameters, is presented in [A.6](#) of Reliability Toolkit, [Appendix A](#).

Transient loads, such as live, snow, wind, earthquake, and flood-induced loads, have event and extreme value distributions. The event distribution quantifies the magnitude of the typical load: for example, the average daily flow in a river. The extreme value distribution quantifies the maximum event that is expected to occur over a specific reference period: for example, the maximum flow that would be expected, still averaged over a 1-day period, over an entire year, or over a

50-year design life. As shown in Fig. 4.2, normal distributions do not accurately represent extreme value distributions, although it is often possible to create an equivalent normal or lognormal distribution that accurately simulates the upper tail of an extreme value distribution. Extreme value distributions have higher mean values and smaller coefficients of variation than their associated event distributions.

The mathematics underlying analysis using extreme value distributions are usually based on the assumption of independent, identically distributed events. This implies, for example, that the maximum extreme load observed this year is independent of the maximum observed in previous years, and that the probability of observing a particular extreme load magnitude remains constant with time. These assumptions do not apply to dead load, for example, because the dead load acting now is usually highly correlated to the dead load that was acting 1 or 2 years ago. These assumptions do not apply to loads that may become more or less severe, and more variable, with the passage of time.

It is common to quantify extreme loads using specified return periods. The return period T is defined as the average number of years to the first occurrence of an event with a specified magnitude. The word “average” is important because it is possible, though unlikely, that the return period event will occur in successive years. Similarly, it is possible, though unlikely, that the next return period event will occur exactly T years after the previous one.

If the events are independent, identically distributed random variables, the probability of occurrence, p , of a return period event in a given year is, by definition, $p = 1/T$. The probability P of equaling or exceeding the return period event in an interval of N years is

$$P = 1 - (1 - 1/T)^N \quad (4.6)$$

4.6.1 Example calculation—What is the probability of a 50-year flood being equaled or exceeded in 1 year, 5 years, 10 years, 25 years, 50 years, 100 years, and 150 years?

These probabilities are readily computed using Eq. (4.6) with $T = 50$ years and $N = 1, 5, 10 \dots 150$ years. For example, for $N = 1$

$$P = 1 - (1 - 1/50)^1 = 1/50 = 0.02 \text{ or } 2\%$$

Table 4.6.1 summarizes the other probability values. There is a 64% probability that the 50-year flood will be equaled or exceeded in 50 years. There is a 4.8% probability that the 50-year flood will not be exceeded in 150 years, which puts into perspective the effectiveness of assessing structural reliability based on the number of years of service of a structure. These values are only accurate if it can be asserted that extreme floods are identically distributed, which implies that

Table 4.6.1—Probability of exceeding 50-year flood in interval N

Interval N , years	1	5	10	25	50	100	150
Probability P	0.020	0.098	0.183	0.397	0.636	0.867	0.952