Statistics Review

for the De Veaux/Velleman/Bock Series

Data

Categorical data are information about characteristics or qualities falling into different classifications.

Quantitative data are information about a quantity or measurement (with units).

The W's describe the data's context—*Who, What, Why, When, Where*, and *hoW*.

Displaying Categorical Data

A **bar chart** displays the distribution of a categorical variable, showing the count or percentage of values in each category.

A **pie chart** displays the distribution of a categorical variable by slicing a circle into pieces whose areas are proportional to the fraction of values in each category.

A **contingency table** displays the distribution of observations categorized on two variables.

Describing Categorical Data

A **distribution** shows counts or percentages of observations in each category.

A **conditional distribution** shows the distribution of one variable within a single category of another variable.

Independence exists when the distribution of one variable is the same in all categories of another variable; if the distribution depends on the category, we say there's an **association**.

Displaying Quantitative Data

A **histogram** displays the distribution of a quantitative variable in bars showing counts or percentages of observations falling in each interval.

A **stem-and-leaf display** records the actual data values falling in each interval by splitting the data into a stem (the tens digit, say) and a leaf (the units digit).

A **dotplot** displays dots (instead of bars or digits) for data values in each interval.

A **boxplot** displays a box spanning the middle 50% of the data (extending from the first to third quartile and showing the median), with whiskers extending to the lowest and highest nonoutlier data values, and outliers plotted.

Describing Quantitative Data

Always mention **shape** (*modes symmetry*), **center**, **spread**, and **unusual features** (*gaps, clusters*, and *outliers*).

Quantitative Data Statistics

Minimum is the smallest data value.

Maximum is the largest data value.

Range: *range* = *maximum* − *minimum*

The **median** is the middle value (half the data are larger, half smaller).

The **quartiles** divide each half of the data in half; 25% of the data are smaller than Q_1 (the first quartile) and 25% larger than Q_3.

Interquartile range: $IQR = Q_3 - Q_1$

Outlier guideline: Data values that lie more than 1.5 IQRs below Q_1 or above Q_3 may be outliers.

5-number summary: *min*, Q_1, *median*, Q_3, *max*

Mean: $\bar{y} = \dfrac{\sum y}{n}$

Standard deviation: $s = \sqrt{\dfrac{\sum(y-\bar{y})^2}{n-1}}$

Two Quantitative Variables

A **scatterplot** displays points corresponding to cases measured on two variables.

The **direction** is *positive* if higher values of one variable are generally associated with higher values of the other and *negative* if higher values of one variable are generally associated with lower values of the other.

Form: *linear* or *curved*

Strength: The less scatter, the stronger the association.

Unusual features: Look for *clusters, outliers*, and *influential points*.

Correlation: r is a number between −1 and +1 describing the direction and strength of a linear relationship between two quantitative variables.

$$r = \frac{\sum z_x z_y}{n-1}$$

Straight Enough Condition: If the pattern in the scatterplot looks reasonably straight, it's okay to fit a linear model.

The **regression line** is a model that predicts a value of y for each x; $\hat{y} = b_0 + b_1 x$, where $b_1 = \dfrac{r s_y}{s_x}$ and $b_0 = \bar{y} - b_1\bar{x}$; the line passes through $(\bar{x}, \bar{y})$.

more

ISBN-13: 978-0-321-82626-8
ISBN-10: 0-321-82626-4
90000
EAN
9 780321 826268

Two Quantitative Variables (continued)

Residual: $e = y - \hat{y}$, the difference between the actual value of y and the value predicted by the model.

Least squares: The regression line minimizes the sum of the squared residuals.

Slope: The slope models the relationship as y-units per x-unit.

Intercept: The y-intercept is the starting value (the value of $\hat{y}$ predicted when $x = 0$).

***R*-squared:** R^2 is the fraction of the variability in y explained by the regression model.

Modeling Wisdom

Residual plots appear as randomly scattered points when the model is appropriate.

Influential points distort the model; if you are suspicious, try creating regression models both with and without them.

Cause and effect: A strong association is *not* evidence of causation.

Subsets: If a scatterplot shows distinct groups, it may be better to fit a model to each one separately.

Curvature: If the relationship is curved, **re-express** one or both variables to straighten the relationship. Possible approaches include the **Ladder of Powers** (re-express y as y^2, $\sqrt{y}$, $\log y$, $\dfrac{-1}{\sqrt{y}}$, $\dfrac{-1}{y}$, etc.) or use $\log y$ and $\log x$.

Simulations

Steps in a simulation: To investigate the distribution of outcomes for a situation of interest, create a simulation model based on random numbers.

1. **Model a component:** Explain how you will interpret random numbers to represent the most basic event of interest.
2. **Simulate a trial:** Explain how you will use random numbers to model one outcome.
3. Define your **response variable**.
4. **Run many trials**, recording the outcome for each.
5. **Analyze the response variable** by graphing the data and calculating summary statistics.
6. **State a conclusion** in the context of the original question.

Sampling

A **sample** is a subset of a **population** for which data are collected and analyzed in an effort to learn about unknown (unknowable) properties of the population. We use **sample statistics** to estimate **population parameters**.

Property	Statistic	Parameter
Proportion	$\hat{p}$	p
Mean	$\bar{y}$	μ
Standard deviation	s	σ
Slope	b_1	β_1

more

Sampling (continued)

Sampling error is sample-to-sample variation in a statistic.

Bias is found in sampling methods that systematically misrepresent characteristics of the population.

- **Undercoverage** limits (or omits) some subpopulation.
- **Voluntary response** allows individuals to self-select their participation.
- **Nonresponse bias** occurs when many of those sampled elect not to participate.
- **Response bias** influences people's answers.

Random sampling gives each member of the population the same chance of being selected.

- In a **simple random sample**, each subset of size n is equally likely to be selected.
- A **stratified sample** draws random samples from each of several homogeneous subpopulations (strata).
- A **cluster sample** randomly selects entire heterogeneous subpopulations (clusters) from among many.
- A **systematic sample** selects (for example) every 12th individual from a list of the population starting from a randomly determined case.

Observational Studies

A **retrospective** study collects information looking into the past; a **prospective** study follows subjects over time.

Observational studies can spot associations between variables but can neither reach conclusions about populations nor establish cause and effect. A **lurking variable** that influences both x and y can make it appear that x causes y.

Experiments

An experiment applies treatments to randomly assigned subjects to observe the response.

A **factor** is a variable manipulated by the experimenter, applied at several different **levels**.

A **treatment** is the combination of factor levels applied to a subject.

The **response variable** is the (usually quantitative) outcome we measure to compare effects of the treatments.

A **control group** receives no treatment (or a null treatment) to provide a baseline for purposes of comparison.

Principles of Design:

- **Control** known sources of variability whenever possible.
- **Randomize** subjects to treatments to balance unknown sources of variability (subjects needn't be a random sample).
- **Replicate** each treatment on many subjects.
- **Block** subjects with respect to preexisting sources of variability we can't control.

Blinding is keeping people involved with the experiment unaware of treatment assignments, both (1) during the experiment (subjects and others in contact with them, often accomplished with **placebos**) and (2) during evaluation of the response. An experiment is **double blind** when both classes are kept unaware.

A **confounding variable** is a variable that influences the response variable in ways that we can't separate from the effects of the experimental factor.

Probability

Probability is the long-run frequency of an event's occurrence; $0 \le P(A) \le 1$.

A sample space is the set of all possible outcomes; $P(S) = 1$.

Complement Rule: $P(A^C) = 1 - P(A)$

Addition Rule: $P(\mathbf{A} \text{ OR } \mathbf{B}) = P(\mathbf{A}) + P(\mathbf{B}) - P(\mathbf{A} \text{ AND } \mathbf{B})$

Multiplication Rule: $P(\mathbf{A} \text{ AND } \mathbf{B}) = P(\mathbf{A}) \times P(\mathbf{B} \mid \mathbf{A})$

Conditional probability: $P(\mathbf{B} \mid \mathbf{A}) = \dfrac{P(\mathbf{A} \text{ AND } \mathbf{B})}{P(\mathbf{A})}$

Disjoint (mutually exclusive) events cannot both happen:

$$P(\mathbf{A} \text{ AND } \mathbf{B}) = 0$$

Independent events: The occurrence of one event has no impact on the probability of the other: $P(\mathbf{A} \mid \mathbf{B}) = P(\mathbf{A})$

For **random variables:**

$$\mu = E(X) = \sum(x \times P(x))$$
$$\sigma^2 = Var(X) = \sum((x-\mu)^2 P(x))$$
$$E(X + c) = E(X) + c$$
$$Var(X + c) = Var(X)$$
$$E(aX) = aE(X)$$
$$Var(aX) = a^2 Var(X)$$
$$E(X \pm Y) = E(X) \pm E(Y)$$

Pythagorean Theorem of Statistics:
If random variables X and Y are independent, then

$$Var(X \pm Y) = Var(X) + Var(Y)$$
$$SD(X \pm Y) = \sqrt{SD^2(X) + SD^2(Y)}.$$

Normal model:

A Normal model, $N(\mu, \sigma)$, is unimodal, symmetric, and bell-shaped and is specified by its mean, μ, and standard deviation, σ.

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

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

99.7% lie within $\mu \pm 3\sigma$.

Bernoulli trials:

- two outcomes (success, failure)
- known probability of success p
- trials are independent

Geometric model:
X = number of Bernoulli trials until the first success

$$P(x) = q^{x-1}p$$
$$E(X) = \frac{1}{p}$$

Binomial model:
X = number of successes in n Bernoulli trials

$$P(x) = \binom{n}{x} p^x q^{n-x}$$
$$E(X) = np \quad SD(X) = \sqrt{npq}$$

Normal approximation: If we expect at least 10 successes and 10 failures, then binomial probabilities may be approximated using the Normal model $N(pq, \sqrt{npq})$.

Sampling Distribution Models

For a sample proportion: Provided that the sampled values are independent and the sample size is large enough, the sampling distribution of $\hat{p}$ can be modeled by a Normal model with

$$\mu(\hat{p}) = p \text{ and } SD(\hat{p}) = \sqrt{\frac{pq}{n}}.$$

For a sample mean: The Central Limit Theorem If a random sample of size n is drawn from a population with mean μ and standard deviation σ, then as n increases the sampling distribution of the sample mean, $\bar{y}$, approaches the Normal model $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$, regardless of the shape of the population.

Confidence Intervals

If the appropriate assumptions and conditions are met, we can have a specified level of confidence that the interval

estimate ± (*critical value*) × *SE*(*estimate*)

captures the value of a population parameter.

Hypothesis Tests

The four steps:

Hypotheses: Write a null hypothesis for the value of the population parameter and specify the alternative hypothesis (upper tail, lower tail, or two-tailed).

Model: Check assumptions and conditions, then specify the type of test and the sampling model.

Mechanics: Calculate the test statistic and find the P-value.

Conclusion: Link the P-value to your decision (reject or fail to reject H_0) and state your conclusion in the proper context.

The **null hypothesis** (H_0) specifies a parameter and a hypothesized value for that parameter.

The **alternative hypothesis** (H_A) is a statement indicating what values of the parameter are of interest (different from, smaller than, or larger than that specified in H_0).

The **P-value** is the probability that results at least as extreme as those we observed could have occurred if the null hypothesis were true.

Type I error is rejecting the null hypothesis when it is true.

Type II error is failing to reject the null hypothesis when it is false.

Power is the probability the test rejects a false null hypothesis.

Effect size is the difference between the hypothesized value of the parameter and its true value.

Assumptions for Inference — And the Conditions That Support or Override Them

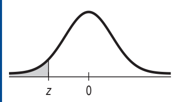
Assumptions for Inference	And the Conditions That Support or Override Them
PROPORTIONS (z)	
• **One sample**	
1. Individuals are independent.	1. SRS and $n < 10\%$ of the population.
2. Sample is sufficiently large.	2. Successes and failures each ≥ 10.
• **Two groups**	
1. Groups are independent.	1. (Think about how the data were collected.)
2. Data in each group are independent.	2. Both are SRSs and $n < 10\%$ of populations OR random allocation.
3. Both groups are sufficiently large.	3. Successes and failures each ≥ 10 for both groups.
MEANS (t)	
• **One Sample** ($df = n - 1$)	
1. Individuals are independent.	1. SRS and $n < 10\%$ of the population.
2. Population has a Normal model.	2. Histogram is unimodal and symmetric.*
• **Matched pairs** ($df = n - 1$)	
1. Data are matched.	1. (Think about the design.)
2. Individuals are independent.	2. SRS and $n < 10\%$ OR random allocation.
3. Population of differences is Normal.	3. Histogram of differences is unimodal and symmetric.*
• **Two independent samples** (df from technology)	
1. Groups are independent.	1. (Think about the design.)
2. Data in each group are independent.	2. SRSs and $n < 10\%$ OR random allocation.
3. Both populations are Normal.	3. Both histograms are unimodal and symmetric.*
DISTRIBUTIONS/ASSOCIATION (χ^2)	
• **Goodness-of-fit** ($df = \#$ of cells $- 1$; one variable, one sample compared with population model)	
1. Data are counts.	1. (Are they?)
2. Data in sample are independent.	2. SRS and $n < 10\%$ of the population.
3. Sample is sufficiently large.	3. All expected counts ≥ 5.
• **Homogeneity** [$df = (r-1)(c-1)$; many groups compared on one variable]	
1. Data are counts.	1. (Are they?)
2. Data in groups are independent.	2. SRSs and $n < 10\%$ OR random allocation.
3. Groups are sufficiently large.	3. All expected counts ≥ 5.
• **Independence** [$df = (r-1)(c-1)$; sample from one population classified on two variables]	
1. Data are counts.	1. (Are they?)
2. Data are independent.	2. SRSs and $n < 10\%$ of the population.
3. Sample is sufficiently large.	3. All expected counts ≥ 5.
REGRESSION (t, $df = n - 2$)	
• **Association** of each *quantitative variable* ($\beta = 0$?)	
1. Form of relationship is linear.	1. Scatterplot looks approximately linear.
2. Errors are independent.	2. No apparent pattern in residuals plot.
3. Variability of errors is constant.	3. Residuals plot has consistent spread.
4. Errors have a Normal model.	4. Histogram of residuals is approximately unimodal and symmetric, or Normal probability plot reasonably straight.*
	(*less critical as n increases)

Quick Guide to Inference

Think			Show			
Inference about?	**One group or two?**	**Procedure**	**Model**	**Parameter**	**Estimate**	**SE**
PROPORTIONS	One sample	1-Proportion z-Interval	z	p	$\hat{p}$	$\sqrt{\frac{\hat{p}\hat{q}}{n}}$
		1-Proportion z-Test				$\sqrt{\frac{p_0 q_0}{n}}$
	Two independent groups	2-Proportion z-Interval	z	$p_1 - p_2$	$\hat{p}_1 - \hat{p}_2$	$\sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$
		2-Proportion z-Test				$\sqrt{\frac{\hat{p}\hat{q}}{n_1} + \frac{\hat{p}\hat{q}}{n_2}}$, $\hat{p} = \frac{y_1 + y_2}{n_1 + n_2}$
MEANS	One sample	t-Interval t-Test	t $df = n - 1$	μ	$\bar{y}$	$\frac{s}{\sqrt{n}}$
	Two independent groups	2-Sample t-Test 2-Sample t-Interval	t df from technology	$\mu_1 - \mu_2$	$\bar{y}_1 - \bar{y}_2$	$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
	Matched pairs	Paired t-Test Paired t-Interval	t $df = n - 1$	μ_d	$\bar{d}$	$\frac{s_d}{\sqrt{n}}$
DISTRIBUTIONS (one categorical variable)	One sample	Goodness-of-Fit	χ^2 df = cells $- 1$			$\sum \frac{(Obs - Exp)^2}{Exp}$
	Many independent groups	Homogeneity χ^2 Test	χ^2 $df = (r-1)(c-1)$			
INDEPENDENCE (two categorical variables)	One sample	Independence χ^2 Test				
ASSOCIATION (two quantitative variables)	One sample	Linear Regression t-Test or Confidence Interval for β	t $df = n - 2$	β_1	b_1	$\frac{s_e}{s_x\sqrt{n-1}}$ (compute with technology)
		*Confidence Interval for μ_ν		μ_ν	$\hat{y}_\nu$	$\sqrt{SE^2(b_1) \times (x_\nu - \bar{x})^2 + \frac{s_e^2}{n}}$
		*Prediction Interval for y_ν		y_ν	$\hat{y}_\nu$	$\sqrt{SE^2(b_1) \times (x_\nu - \bar{x})^2 + \frac{s_e^2}{n} + s_e^2}$

Inference about?	**One group or two?**	**Procedure**	**Model**	**Parameter**	**Estimate**	**SE**	**Chapter**

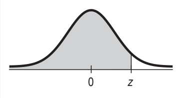
Table Z
Areas under the standard normal curve



Second decimal place in z										
0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00	z
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	−3.8
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	−3.7
0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0002	0.0002	−3.6
0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	−3.5
0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	−3.4
0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005	−3.3
0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007	−3.2
0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010	−3.1
0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013	−3.0
0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019	−2.9
0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026	−2.8
0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035	−2.7
0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047	−2.6
0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062	−2.5
0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082	−2.4
0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107	−2.3
0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139	−2.2
0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179	−2.1
0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228	−2.0
0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287	−1.9
0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359	−1.8
0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446	−1.7
0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548	−1.6
0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668	−1.5
0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808	−1.4
0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968	−1.3
0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0.1151	−1.2
0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357	−1.1
0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587	−1.0
0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841	−0.9
0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119	−0.8
0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420	−0.7
0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743	−0.6
0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085	−0.5
0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446	−0.4
0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821	−0.3
0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207	−0.2
0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602	−0.1
0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000	−0.0

For $z \leq -3.90$, the areas are 0.0000 to four decimal places.

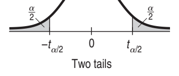
Table Z (cont.)
Areas under the standard normal curve

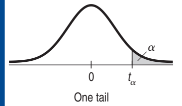


	Second decimal place in z									
z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999

For $z \geq 3.90$, the areas are 1.0000 to four decimal places.

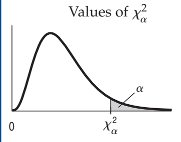
Table T
Values of t_α





Two tail probability / One tail probability	0.20 / 0.10	0.10 / 0.05	0.05 / 0.025	0.02 / 0.01	0.01 / 0.005	
df						df
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
10	1.372	1.812	2.228	2.764	3.169	10
11	1.363	1.796	2.201	2.718	3.106	11
12	1.356	1.782	2.179	2.681	3.055	12
13	1.350	1.771	2.160	2.650	3.012	13
14	1.345	1.761	2.145	2.624	2.977	14
15	1.341	1.753	2.131	2.602	2.947	15
16	1.337	1.746	2.120	2.583	2.921	16
17	1.333	1.740	2.110	2.567	2.898	17
18	1.330	1.734	2.101	2.552	2.878	18
19	1.328	1.729	2.093	2.539	2.861	19
20	1.325	1.725	2.086	2.528	2.845	20
21	1.323	1.721	2.080	2.518	2.831	21
22	1.321	1.717	2.074	2.508	2.819	22
23	1.319	1.714	2.069	2.500	2.807	23
24	1.318	1.711	2.064	2.492	2.797	24
25	1.316	1.708	2.060	2.485	2.787	25
26	1.315	1.706	2.056	2.479	2.779	26
27	1.314	1.703	2.052	2.473	2.771	27
28	1.313	1.701	2.048	2.467	2.763	28
29	1.311	1.699	2.045	2.462	2.756	29
30	1.310	1.697	2.042	2.457	2.750	30
32	1.309	1.694	2.037	2.449	2.738	32
35	1.306	1.690	2.030	2.438	2.725	35
40	1.303	1.684	2.021	2.423	2.704	40
45	1.301	1.679	2.014	2.412	2.690	45
50	1.299	1.676	2.009	2.403	2.678	50
60	1.296	1.671	2.000	2.390	2.660	60
75	1.293	1.665	1.992	2.377	2.643	75
100	1.290	1.660	1.984	2.364	2.626	100
120	1.289	1.658	1.980	2.358	2.617	120
140	1.288	1.656	1.977	2.353	2.611	140
180	1.286	1.653	1.973	2.347	2.603	180
250	1.285	1.651	1.969	2.341	2.596	250
400	1.284	1.649	1.966	2.336	2.588	400
1000	1.282	1.646	1.962	2.330	2.581	1000
∞	1.282	1.645	1.960	2.326	2.576	∞
Confidence levels	80%	90%	95%	98%	99%	

Table χ
Values of χ^2_α



Right tail probability	0.10	0.05	0.025	0.01	0.005
df					
1	2.706	3.841	5.024	6.635	7.879
2	4.605	5.991	7.378	9.210	10.597
3	6.251	7.815	9.348	11.345	12.838
4	7.779	9.488	11.143	13.277	14.860
5	9.236	11.070	12.833	15.086	16.750
6	10.645	12.592	14.449	16.812	18.548
7	12.017	14.067	16.013	18.475	20.278
8	13.362	15.507	17.535	20.090	21.955
9	14.684	16.919	19.023	21.666	23.589
10	15.987	18.307	20.483	23.209	25.188
11	17.275	19.675	21.920	24.725	26.757
12	18.549	21.026	23.337	26.217	28.300
13	19.812	22.362	24.736	27.688	29.819
14	21.064	23.685	26.119	29.141	31.319
15	22.307	24.996	27.488	30.578	32.801
16	23.542	26.296	28.845	32.000	34.267
17	24.769	27.587	30.191	33.409	35.718
18	25.989	28.869	31.526	34.805	37.156
19	27.204	30.143	32.852	36.191	38.582
20	28.412	31.410	34.170	37.566	39.997
21	29.615	32.671	35.479	38.932	41.401
22	30.813	33.924	36.781	40.290	42.796
23	32.007	35.172	38.076	41.638	44.181
24	33.196	36.415	39.364	42.980	45.559
25	34.382	37.653	40.647	44.314	46.928
26	35.563	38.885	41.923	45.642	48.290
27	36.741	40.113	43.195	46.963	49.645
28	37.916	41.337	44.461	48.278	50.994
29	39.087	42.557	45.722	59.588	52.336
30	40.256	43.773	46.979	50.892	53.672
40	51.805	55.759	59.342	63.691	66.767
50	63.167	67.505	71.420	76.154	79.490
60	74.397	79.082	83.298	88.381	91.955
70	85.527	90.531	95.023	100.424	104.213
80	96.578	101.879	106.628	112.328	116.320
90	107.565	113.145	118.135	124.115	128.296
100	118.499	124.343	129.563	135.811	140.177