Appendix A
Writing Lab Reports

To prepare your lab report, distill your investigations into a central argument. Then organize your findings into units of evidence for your argument. A good lab report is more than a compilation of your statistical analyses. A good report should be well organized, and it should demonstrate clear and sound reasoning, contain easily interpreted data displays, and use good grammar.

Organization[1]

Format your report into recognizable sections to clarify the structure of the paper. Two levels of headings are usually helpful to provide a general outline. A paper without section headings drags and is difficult to follow.

The choice of sections should match the reader's expectations. For example, a research article is generally divided into sections labeled: Introduction, Methodology, Results, and Discussion. When the sections are jumbled, such as when discussion or experimental detail is found in the statement of the results, the readers may become confused.

Use the Introduction to state the problem you are addressing and your findings. Without giving away all your points, let the reader know where your paper is headed.

---

[1]The first three paragraphs of this section are summarized from, S. Tollefson, "Encouraging Student Writing," S. Tollefson, University of California, p. 24.

- Catch the reader's attention. Start with an example, a quotation, a statistic, a question, or a complaint and use it as a theme that you refer to throughout the paper.
- The Introduction sets the tone for your report. Explain why the problem you are addressing is important. Appear to be interested in the topic.
- Break up a long Introduction into several paragraphs. One huge paragraph at the outset of a paper can put readers off.
- Avoid such phrases as "I will discuss" or "this report will examine." Better to just dive right in.

Use the Methodology section to describe your data and how they were collected. This information helps the reader assess the appropriateness of your analysis and the significance of your findings.

- Describe the subject(s) under study. Be as specific as possible. Make clear who was included in the study and who was not.
- Outline the procedures used for collecting the data. For example, if the data are from a sample survey, then the reader may need to know the sampling method, the interview process, and the exact wording of the questions asked. Also address any problems with the data such as nonresponse.
- Explain how the variables measured can be used to address the scientific question of interest. Clearly distinguish between the main outcome of the study and auxiliary information. Be sure to provide the units in which the responses were measured.

Use the Results section to present your findings. Limit the presentation to those results that are most relevant to your argument and most understandable to the reader.

- Be parsimonious in your use of supporting tables and graphs. Too much extraneous information overloads the reader and obscures the importance of your main thesis. The reader is often willing to accept a brief statement summarizing your additional findings, especially if the material presented is well displayed and to the point. When preparing your data displays, follow the guidelines that appear later in this appendix. Each display must be discussed in the prose.
- Limit the use of statistical jargon. Save the most technical material for an Appendix where you show the advanced reader your more sophisticated ideas and more complicated calculations.
- Include in your report any findings that point to a potential shortcoming in your argument. These problems should be considered in the Discussion section.
- If you include a figure from another paper, cite the original source in your figure caption.

Use the Discussion section to pull together your results in defense of your main thesis.

- Be honest. Address the limitations of your findings. Discuss, if possible, how your results can be generalized.

- Be careful not to overstate the importance of your findings. With statistical evidence, we can rarely prove a conjecture or definitively answer a question. More often than not, the analysis provides support for or against a theory, and it is your job to assess the strength of the evidence presented.
- Relate your results to the rest of the scientific literature. Remember to give credit to the ideas of others. Consider the following questions:
  - — Do your results confirm earlier findings or contradict them?
  - — What additional information does your study provide over past studies?
  - — What are the unique aspects of your analysis?
  - — If you could continue research into the area, what would you suggest for the next step?

Clarity and Structure of Prose[2]

Information in a passage of text is interpreted more easily and more uniformly if it is placed where readers expect to find it. Readers naturally emphasize the material that arrives at the end of a sentence. When the writer puts the emphatic material at the beginning or middle of a sentence, the reader is highly likely to emphasize the wrong material and to incorrectly interpret the message of the sentence. Readers also expect the material at the beginning of the sentence to provide them a link to previous material and a context for upcoming material. When old information consistently arrives at the beginning of the sentence, it helps readers to construct the logical flow of the argument.

Observe the following structural principles:

- Follow a grammatical subject as soon as possible with its verb.
- Place at the end of the sentence the new information you want the reader to emphasize.
- Place the person or thing whose story a sentence is telling at the beginning of the sentence.
- Place appropriate old information (material already stated in the discourse) at the beginning of the sentence for linkage backward and contextualization forward.
- Provide context for your reader before asking that reader to consider anything new.
- Try to ensure that the relative emphases of the substance coincide with the relative expectations raised by the structure.

Here is an example from Gopen and Swan of scientific prose that begins sentences with new information and ends with old information. After reading the paragraph, we have no clear sense of where we have been or where we are going.

---

[2]The material in this section is summarized from G.D. Gopen and J.A. Swan, "The Science of Scientific Writing," *Am. Sci.*, **78**:550–558.

> Large earthquakes along a given fault segment do not occur at random intervals because it takes time to accumulate the strain energy for the rupture. The rates at which tectonic plates move and accumulate strain at their boundaries are approximately uniform. Therefore, in first approximation, one may expect that large ruptures of the same fault segment will occur at approximately constant time intervals. If subsequent main shocks have different amounts of slip across the fault, then the recurrence time may vary, and the basic idea of periodic mainshocks must be modified. For great plate boundary ruptures the length and slip often vary by a factor of 2. Along the southern segment of the San Andreas fault the recurrence interval is 145 years with variations of several decades. The smaller the standard deviation of the average recurrence interval, the more specific could be the long term prediction of a future mainshock.

Gopen and Swan revised the paragraph to abide by the structural principles outlined above. The phrases in square brackets are suggestions for connections between sentences. These connections were left unarticulated in the original paragraph; they point out the problems that had existed with the logical flow of the argument.

> Large earthquakes along a given fault segment do not occur at random intervals because it takes time to accumulate the strain energy for the rupture. The rates at which tectonic plates move and accumulate strain at their boundaries are approximately uniform. Therefore, nearly constant time intervals (at first approximation) would be expected between large ruptures of the same fault segment. [However?], the recurrence time may vary; the basic idea of periodic mainshocks may need to be modified if subsequent main shocks have different amounts of slip across the fault. [Indeed?], the length and slip of great plate boundary ruptures often vary by a factor of 2. [For example?], the recurrence intervals along the southern segment of the San Andreas fault is 145 years with variations of several decades. The smaller the standard deviation of the average recurrence interval, the more specific could be the long term prediction of a future mainshock.

Data Displays[3]

The aim of good data graphics is to display data accurately and clearly, and the rules for good data display are quite simple. Examine data carefully enough to know what they have to say, and then let them say it with a minimum amount of adornment. Do this while following reasonable regularity practices in the depiction of scale, and label clearly and fully.

[3]The material in this section is summarized from "How to Display Data Badly," H. Wainer, *The American Statistician* **38**:137–147, 1984.

The following list provides guidelines for how to make good data displays.

- *Density*– Holding clarity and accuracy constant, the more information displayed the better. When a graph contains little information, the plot looks empty and raises suspicions that there is nothing to be communicated. However, avoid adding to the displays extraneous graphics such as three-dimensional bars, stripes, and logos. Chart junk does not increase the quantity of information conveyed in the display; it only hides it.
- *Scale*–
  - Graph data in context; show the scale of your axes.
  - Choose a scale that illuminates the variation in the data.
  - Do not change scale in mid-axis.
  - If two plots are to be compared, make their scales the same.
- *Labels*– Captions, titles, labels, and legends must be legible, complete, accurate, and clear.
- *Precision*– Too many decimal places can make a table hard to understand. The precision of the data should dictate the precision reported. For example, if weight is reported to the nearest 5 pounds then a table presenting average weights should not be reported to the nearest 1/100 of a pound.
- *Dimensions*– If the data are one-dimensional, then use a visual metaphor that is one-dimensional. Increasing the number of dimensions can make a graph more confusing. Additional dimensions can cause ambiguity: is it length, area, or volume that is being compared?
- *Color*– Adding color to a graph is similar to adding an extra dimension to the graph. The extra dimension should convey additional information. Using color in a graph can make us think that we are communicating more than we are.
- *Order*– Sometimes the data that are to be displayed have one important aspect and others that are trivial. Choose a display that makes it easy to make the comparison of greatest interest. For example: (a) ordering graphs and tables alphabetically can obscure structure in the data that would have been obvious had the display been ordered by some aspect of the data; (b) Stacking information graphically indicates the total but can obscure the changes in individual components, especially if one component both dominates and fluctuates greatly; (c) comparisons are most easily made by placing information all on one plot.

Grammar[4]

Bad grammar distracts the reader from what you are trying to say. Always assume that whoever reads your paper will be evaluating your grammar, sentence structure, and style, as well as content. Content cannot really be divorced from form.

---

[4]The material in this section is reprinted with permission from S. Tollefson, "Encouraging Student Writing," University of California, p. 27.

Examples of grammatical problems:

- *subject–verb or noun–pronoun agreement*
  Theories of cosmology suggests that the universe must have more than three dimensions. (“suggest”)
  Everyone investigating mass extinctions by asteroids may bring their own prejudices to the investigations. (“Scientists investigating ... may bring their...”)
- *faulty comparison—either incomplete or mixing apples and oranges*
  Thin metal strands deposited in a silicon wafer make it better. (“better” than what?)
  The stars in some galaxies are much more densely packed than the Milky Way. (“than those in the Milky Way”)
- *sentence fragment*
  The HTLV virus, nearly unknown six years ago (although some evidence of a longer existence has been found) rapidly becoming the Black Plague of modern times. (add “is” before “rapidly”)
  The virus may be handled in the laboratory. But only with care. (These should be one sentence.)
- *misuse of tenses and confusion of parts of speech*
  Some researchers feel badly about how lab animals are treated. (“feel bad”)
  Ever since Quagmire Chemical Company requested a heat exchanger, we investigated several options. (“have investigated”)
- *idiom—usually a misuse of prepositions*
  The Human Subjects Committee insisted to do it their way (“insisted on doing it...”)
- *modification—usually a word or phrase in the wrong place*
  When applying an electric field, liquid crystal molecules align themselves and scattering of light is reduced. (Dangling—should be “When one applies...”)
  Incinerating industrial wastes can produce compounds toxic to humans such as dioxins. (misplaced—should be “can produce compounds, such as dioxins, toxic to humans.”)
- *parallel structure*
  In gel electrophoresis, large ions move slowly and small ones are traveling more quickly. (“...and small ones travel more quickly”)
  The vaccinia virus may be used to vaccinate against small pox and as a carrier of genetic material from one organism to another. (“...to vaccinate...and to carry”)
- *passive voice—not always bad, but overused*
  Short stature and low IQ can be caused by an extra chromosome. (more often than not, it’s preferable to say “An extra chromosome causes...”)
  Trisomy-21 is still often called “mongolism.” (“People still often call...”)
- *predication—illogical connection among subject/verb/complements*
  Viewing occultations of certain asteroids suggests that they have moons. (it’s not the viewing, but the occultations themselves that suggest this)
  Exposure to intense x-rays is the reason these crystals glow. (exposure itself

is not the reason—what the x-rays do to the structure causes the glow; or you could say "Exposure to intense x-rays causes these crystals to glow.")

- *reference—faulty or vague*
  The deprojector and image stretcher allow us to examine tilted galaxies as if they were facing us head on. It's a great breakthrough. ("Development of these devices is a great...")
- *run-together sentence*
  The meteor impact site in the Indian Ocean is the best possibility at the moment, however, other sites do exist. (semicolon or period needed before "however")
  The layer of semenium is worldwide, it shows only a few gaps. (semicolon or period needed after "worldwide")

Revising and Proofreading

After you have completed a draft, look at the paper again. Learn to see where new material is needed, where material should be deleted, and where reorganization is required. Once you have spent enormous effort on the analysis and writing, proofread your manuscript two or three times. Keep looking for unclear passages, format errors, poor reasoning, etc., right down to the moment you submit the paper.

Use the following questions to appraise your manuscript:

- Is the problem clearly stated?
- Are the statistical statements correct?
- Are the data displays informative?
- Are the conclusions based on sound evidence?
- Are the grammar and sentence structure correct?
- Are the style and tone appropriate for the venue?

Appendix B
Probability

Random Variables

A *discrete* random variable is a random variable that can take on a finite or at most a countably infinite number of different values. If a discrete random variable X takes values $x_1, x_2, \ldots$, its distribution can be described by its *probability mass function*, or *frequency function*,

$$f(x_i) = \mathbb{P}(X = x_i).$$

These satisfy $0 \le f(x_i) \le 1$ and $\sum_i f(x_i) = 1$.

The *distribution function* of any random variable X is defined as

$$F(x) = \mathbb{P}(X \le x).$$

$F(x)$ is a nondecreasing, left-continuous function which can take values between 0 and 1.

The *expected value* of a discrete random variable X with probability mass function f is defined as

$$\mathbb{E}(X) = \sum_i x_i f(x_i).$$

One can also compute the expected value of any function $g(X)$ by

$$\mathbb{E}[g(X)] = \sum_i g(x_i) f(x_i).$$

A *continuous* random variable is a random variable that can take on uncountably many values, for example, all real numbers, or all real numbers in the interval $[-1,$

1]. The analog of the probability mass function for continuous random variables is the *density function* $f(x)$, which is a nonnegative, piecewise continuous function such that $\int_{-\infty}^{\infty} f(x)dx = 1$. For any real numbers $a \leq b$,

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx.$$

It follows that for any single point a, $\mathbb{P}(X = a) = 0$. Additionally, we may write the distribution function as

$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)dt.$$

Conversely, the density may be expressed through the distribution function by

$$f(x) = \frac{d}{dx}F(x).$$

The expected value of a continuous random variable X with density function f is

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x)dx,$$

and for any function $g(X)$,

$$\mathbb{E}\left[g(X)\right] = \int_{-\infty}^{\infty} g(x)f(x)dx.$$

The *joint distribution function* of variables X and Y is

$$F(x, y) = \mathbb{P}(X \leq x, Y \leq y).$$

For two discrete random variables, one can define their joint probability mass function as

$$f(x, y) = \mathbb{P}(X = x, Y = y).$$

Then the *marginal* probability mass functions of X and Y can be computed as, respectively,

$$f_X(x) = \sum_y f(x, y) \quad \text{and} \quad f_Y(y) = \sum_x f(x, y).$$

One can similarly define the joint density function $f(x, y)$ of two continuous random variables X and Y. It is a nonnegative piecewise continuous function of two variables x and y such that $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)dxdy = 1$. The marginal densities of X and Y are then defined by

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy \quad \text{and} \quad f_Y(y) = \int_{-\infty}^{\infty} f(x, y)dx.$$

The conditional distribution of Y given X is a probability mass function (or a density function) written $f(y|x)$ such that

$$f(x, y) = f(y|x)f_X(x).$$

Two random variables X and Y are *independent* if their joint probability mass function or density $f(x, y)$ can be written as

$$f(x, y) = f_X(x) f_Y(y),$$

or equivalently $f(y|x) = f_Y(y)$.

Intuitively, independence means that knowing the value of one variable does not change the distribution of the other variable. Independent variables X and Y have the following important property:

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y),$$

and in general, for any functions g and h,

$$\mathbb{E}\left[g(X)h(Y)\right] = \mathbb{E}\left[g(X)\right]\mathbb{E}\left[h(Y)\right].$$

Properties of Expectation, Variance and Covariance

The *variance* of a random variable X is defined as

$$\text{Var}(X) = \mathbb{E}\left[X - \mathbb{E}(X)\right]^2 = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2.$$

The *standard deviation* of X is defined as

$$\text{SD}(X) = \sqrt{\text{Var(X)}}.$$

The *covariance* of two random variables X and Y is

$$\text{Cov}(X, Y) = \mathbb{E}[X - \mathbb{E}(X)][Y - \mathbb{E}(Y)] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y),$$

and the *correlation coefficient* is

$$\text{corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\text{SD}(X)\text{SD}(Y)}.$$

X and Y are called *uncorrelated* if their correlation coefficient is equal to 0. Independent variables are always uncorrelated, but uncorrelated variables are not necessarily independent.

If a and b are real numbers, and X and Y are random variables, then

1. $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$.
2. $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.
3. $\text{Var}(X) \geq 0$.
4. $\text{Var}(aX + b) = a^2\text{Var}(X)$.
5. $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$.
6. If X and Y are uncorrelated, $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$
7. $-1 \leq \text{corr}(X, Y) \leq 1$, and if $\text{corr}(X, Y) = \pm 1$, then there exist some constants a and b such that $Y = aX + b$.

Examples of Discrete Distributions

1. *Bernoulli* with parameter p, $0 \le p \le 1$.

- X can be thought of as an outcome of a coin toss ($X = 1$ if head; $X = 0$ if tail) for a coin with $\mathbb{P}(\text{head}) = p$.
- X can take values 0 or 1.

$$f(1) = \mathbb{P}(X = 1) = p, \quad f(0) = \mathbb{P}(X = 0) = 1 - p;$$

$$\mathbb{E}(X) = p, \quad \text{Var}(X) = p(1 - p).$$

2. *Binomial* with parameters n and p, n a positive integer and $0 \le p \le 1$.

- X can be thought of as the number of heads in n independent coin tosses for a coin with $\mathbb{P}(\text{head}) = p$.
- X can take values $k = 0, 1, \ldots, n$.

$$f(k) = \binom{n}{k} p^k (1 - p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1 - p)^{n-k}.$$

$$\mathbb{E}(X) = np, \quad \text{Var}(X) = np(1 - p).$$

- If $X_1, \ldots, X_n$ are independent Bernoulli random variables with common parameter p, then $X = X_1 + \ldots + X_n$ is binomial with parameters n and p.
- A useful identity for binomial coefficients:

$$\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x + y)^n.$$

- If we let $x = p$, $y = 1 - p$, we see $\sum_{k=0}^{n} f(k) = 1$. If we let $x = y = 1$, we get the identity

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

3. *Geometric* with parameter p, $0 \le p \le 1$.

- X can be thought of as the number of coin tosses up to and including the first head for a coin with $\mathbb{P}(\text{head}) = p$.
- X can take values $k = 1, 2, \ldots$.

$$f(k) = (1 - p)^{k-1} p.$$

$$\mathbb{E}(X) = \frac{1}{p}, \quad \text{Var}(X) = \frac{1 - p}{p^2}.$$

4. *Negative binomial* with parameters r and p, r a positive integer and $0 \le p \le 1$.

- X can be thought of as the number of coin tosses up to and including the time a head appears for the rth time for a coin with $\mathbb{P}(\text{head}) = p$.

- X can take values $k = r, r+1, \ldots$.

$$f(k) = \binom{k-1}{r-1}(1-p)^{k-r}p^r.$$

$$\mathbb{E}(X) = \frac{r}{p}, \quad \mathrm{Var}(X) = \frac{r(1-p)}{p^2}.$$

- If $X_1, \ldots, X_r$ are independent geometric random variables with parameter p, then $X = X_1 + \ldots + X_r$ is negative binomial with parameters r and p.

5. *Poisson* with parameter $\lambda > 0$.

- X can take values $k = 0, 1, \ldots$.

$$f(k) = \frac{e^{-\lambda}\lambda^k}{k!}.$$

$$\mathbb{E}(X) = \lambda, \quad \mathrm{Var}(X) = \lambda.$$

- If $X_1, \ldots, X_n$ are independent Poisson random variables with parameters $\lambda_1, \ldots, \lambda_n$, then their sum $X = X_1 + \cdots + X_n$ is a Poisson random variable with parameter $\lambda_1 + \cdots + \lambda_n$.

6. *Hypergeometric* with parameters n, M, and N all positive integers.

- Suppose an urn contains N balls: M black and $N - M$ white. If n balls are drawn without replacement, then X denotes the number of black balls among them.
- X can take values $k = \max(0, n - N + M), \ldots, \min(n, M)$.

$$f(k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}},$$

$$\mathbb{E}(X) = n\frac{M}{N}, \quad \mathrm{Var}(X) = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}.$$

- The hypergeometric distribution describes sampling without replacement from a finite population. If we keep n fixed and let M and N go to infinity in such a way that $M/N \to p$, then the hypergeometric probability mass function converges to the binomial probability mass function with parameters n and p. In other words, when the population is very large, there is essentially no difference between sampling with and without replacement.

7. *Multinomial* with parameters $n, p_1, \ldots, p_r$, n a positive integer, $0 \le p_i \le 1$, $\sum_{i=1}^r p_i = 1$.

- Suppose each of n independent trials can result in an outcome of one of r types, with probabilities $p_1, \ldots, p_r$, and let X_i be the total number of outcomes of type i in n trials.

- Each X_i can take values $n_i = 0, 1, \ldots, n$, such that $\sum_i^n X_i = n$.

$$\mathbb{P}(X_1 = n_1, \ldots, X_r = n_r) = \binom{n}{n_1, \ldots, n_r} p_1^{n_1} p_2^{n_2} \ldots p_r^{n_r},$$

provided that $n_1 + n_2 + \ldots + n_r = n$, where

$$\binom{n}{n_1, \ldots, n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}.$$

- The binomial distribution is a special case of multinomial with $r = 2$.

Examples of Continuous Distributions

1. *Uniform* on the interval (0, 1).

$$f(x) = \begin{cases} 1 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases},$$

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } 0 \le x \le 1 \\ 1 & \text{if } x > 1 \end{cases},$$

$$\mathbb{E}(X) = \frac{1}{2}, \quad \text{Var}(X) = \frac{1}{12}.$$

Uniform on (0, 1) can be generalized to uniform on any interval (a, b) for a, b real numbers with $a < b$. If X is uniform on (0, 1), then $X' = a + (b - a)X$ is uniform on (a, b). It is easy to see that X' has density

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

and

$$\mathbb{E}(X') = \frac{a+b}{2}, \quad \text{Var}(X') = \frac{(b-a)^2}{12}.$$

2. *Exponential* with parameter $\lambda > 0$.

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \ge 0 \\ 0 & \text{otherwise} \end{cases},$$

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \ge 0 \\ 0 & \text{otherwise} \end{cases},$$

$$\mathbb{E}(X) = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}.$$

- The Poisson process provides an important connection between exponential and Poisson random variables. Suppose $X_1, X_2, \ldots$ are independent exponentials with parameter λ. Let $S_0 = 0$, and $S_n = X_1 + \cdots + X_n$ for all n. Fix some $t > 0$. Let $N(t)$ be the last index n such that $S_n < t$, i.e. $S_{N(t)} < t \le S_{N(t)+1}$. Then $N(t)$ is a Poisson random variable with parameter λt. Moreover, if we think of $S_1, \ldots, S_n$ as points on a real line, then, conditional on $N(t) = n$, the locations of these n points are independently and uniformly distributed in the interval $(0, t)$.

3. *Gamma* with parameters $\lambda > 0$ and $\alpha > 0$.

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} & \text{if } x \ge 0 \\ 0 & \text{otherwise} \end{cases},$$

where the gamma function $\Gamma(\alpha)$ is defined by $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ and for α a positive integer $\Gamma(\alpha) = (\alpha - 1)!$.

$$\mathbb{E}(X) = \frac{\alpha}{\lambda}, \quad \mathrm{Var}(X) = \frac{\alpha}{\lambda^2}.$$

- Exponential with parameter λ is a special case of Gamma with $\alpha = 1$. Additionally, if $X_1, \ldots, X_n$ are independent exponential random variables with parameter λ, then $X = X_1 + \cdots + X_n$ has a Gamma distribution with parameters λ and n.

4. *Beta* with parameters $a > 0$ and $b > 0$.

$$f(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases},$$

where the beta function $B(a, b)$ is defined by

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

$$\mathbb{E}(X) = \frac{a}{a+b}, \quad \mathrm{Var}(X) = \frac{ab}{(a+b)^2(a+b+1)}.$$

5. *Standard normal*, written $\mathcal{N}(0, 1)$.

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right),$$

$$\mathbb{E}(X) = 0, \quad \mathrm{Var}(X) = 1.$$

6. *Normal* with parameters μ and $\sigma^2 > 0$, written $\mathcal{N}(\mu, \sigma^2)$.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right),$$

$$\mathbb{E}(X) = \mu, \quad \text{Var}(X) = \sigma^2.$$

- If X is $\mathcal{N}(\mu, \sigma^2)$, then $(X - \mu)/\sigma$ is $\mathcal{N}(0, 1)$, a standard normal. If Y is a standard normal, then $\sigma Y + \mu$ is $\mathcal{N}(\mu, \sigma^2)$.
- If X_i's are independent $\mathcal{N}(\mu_i, \sigma_i^2)$, and $a_1, \dots, a_n$ are real numbers, then $X = a_1 X_1 + \cdots + a_n X_n$ is $\mathcal{N}(\sum_{i=1}^n a_i \mu_i, \sum_{i=1}^n a_i^2 \sigma_i^2)$.

7. *Lognormal* with parameters μ and $\sigma^2 > 0$.

$$f(x) = \begin{cases} \dfrac{1}{x\sqrt{2\pi\sigma^2}} \exp\left(-\dfrac{1}{2\sigma^2}(\log y - \mu)^2\right) & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases},$$

$$\mathbb{E}(X) = \exp\left(\mu + \frac{1}{2}\sigma^2\right), \quad \text{Var}(X) = \exp(2\mu + \sigma^2)\left(\exp(\sigma^2) - 1\right).$$

- X is called lognormal if the logarithm of X has a normal distribution with parameters μ and σ^2, or equivalently, if U is $\mathcal{N}(\mu, \sigma^2)$, then $X = e^U$.

8. *Chi-square* distribution (χ_n^2) with parameter n, a positive integer called *degrees of freedom*.

- The easiest way to define χ_n^2 is as follows: let $X_1, \dots, X_n$ be independent normal random variables with mean 0 and variance 1. Then $X_1^2 + X_2^2 + \cdots + X_n^2$ has a χ_n^2 distribution.
- χ_n^2 is a special case of Gamma distribution with $\lambda = 1/2$ and $\alpha = n/2$.

$$\mathbb{E}(\chi_n^2) = n, \quad \text{Var}(\chi_n^2) = 2n.$$

- If $X_1, \dots, X_n$ are a sample from $\mathcal{N}(\mu, \sigma^2)$, with $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i$ the sample mean and $s^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{x})^2$ the sample variance, then

$$\frac{n-1}{\sigma^2}s^2$$

has a χ_{n-1}^2 distribution.

9. *t distribution*, or *Student's distribution* (t_n), with a positive integer parameter n (also called degrees of freedom).

- If Z is $\mathcal{N}(0, 1)$, Y is χ_n^2, and Z and Y are independent, then

$$\frac{Z}{\sqrt{Y/n}}$$

has a t_n distribution.
- When n is large, the t_n distribution becomes very similar to the standard normal distribution.

$$\mathbb{E}(t_n) = 0 \ \text{ for } n > 1; \quad \text{Var}(t_n) = \frac{n}{n-2} \ \text{ for } n > 2;$$

otherwise $\mathbb{E}(t_n)$ and $\mathrm{Var}(t_n)$ are not defined.

- If $x_1, \ldots, x_n$ are a sample from $\mathcal{N}(\mu, \sigma^2)$, with $\bar{x}$ the sample mean and s^2 the sample variance, then

$$\frac{\bar{x} - \mu}{s/\sqrt{n}}$$

has a t_{n-1} distribution. Here we are using another important property of normal sample statistics: $\bar{x}$ and s^2 are independent.

10. *F distribution* (F_{n_1,n_2}) with n_1 and n_2 degrees of freedom.

- If Y_1 and Y_2 are independent $\chi^2_{n_1}$ and $\chi^2_{n_2}$ random variables, then

$$\frac{Y_1/n_1}{Y_2/n_2}$$

has an F distribution with n_1 and n_2 degrees of freedom.

$$\mathbb{E}(F_{n_1,n_2}) = \frac{n_2}{n_2 - 2} \text{ for } n_2 > 2.$$

11. *Standard bivariate normal*, written $\mathcal{N}_2(0, 0, 1, 1, \rho)$, where $|\rho| \leq 1$.

$$f(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right).$$

12. *Bivariate normal* with parameters $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$ and ρ: μ_X and μ_Y are real numbers; $\sigma_X^2, \sigma_Y^2 > 0$; $|\rho| \leq 1$. Written $\mathcal{N}_2(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$.

- If (X, Y) has a $\mathcal{N}_2(0, 0, 1, 1, \rho)$ standard bivariate normal distribution, then $(\sigma_X X + \mu_X, \sigma_Y Y + \mu_Y)$ has a bivariate normal distribution $\mathcal{N}_2(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$. Similarly, if $(\tilde{X}, \tilde{Y})$ is bivariate normal distribution $\mathcal{N}_2(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho)$, then $((\tilde{X} - \mu_X)/\sigma_X, (\tilde{Y} - \mu_Y)/\sigma_Y)$ is $\mathcal{N}_2(0, 0, 1, 1, \rho)$.
- Marginal distributions of X and Y are $\mathcal{N}(\mu_X, \sigma_X^2)$ and $\mathcal{N}(\mu_Y, \sigma_Y^2)$, respectively, and $\rho = \mathrm{corr}(X, Y)$. If $\rho = 0$, then X and Y are independent (note that this property does not hold in general).
- The conditional distribution of Y given $X = x$ is normal with mean $\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X)$ and variance $\sigma_Y^2(1 - \rho^2)$.
- Note that if two variables have univariate normal distributions, their joint distribution is not necessarily bivariate normal.

Limit Theorems

1. *Law of Large Numbers:* Let $X_1, X_2, \ldots$ be a sequence of independent random variables with mean μ and variance σ^2. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. Then for any $t > 0$,

$$\mathbb{P}(|\bar{X}_n - \mu| > t) \to 0 \text{ as } n \to \infty.$$

2. *Central Limit Theorem:* Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random variables with mean μ and variance σ^2. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Let $\Phi(z) = \mathbb{P}(\mathcal{N}(0,1) \le z)$ be the distribution function of a standard normal. Then for any z,

$$\left|\mathbb{P}\left(\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \le z\right) - \Phi(z)\right| \to 0 \text{ as n} \to \infty.$$

This mode of convergence is called convergence in distribution.

3. *Another Central Limit Theorem:* Let $x_1, \ldots, x_N$ be real numbers, and fix n, $0 < n < N$. Let S be a simple random sample of size n from $x_1, \ldots, x_N$, taken without replacement. Let $F(t)$ be the distribution function of the sample sum. Then for any t (Höglund [Hog78]),

$$\left|F(t) - \Phi\left(\frac{t - n\bar{x}}{s\sqrt{pq}}\right)\right| \le \frac{C\sum_{k=1}^{N}(x_k - \bar{x})^3}{s^3\sqrt{pq}},$$

where $p = n/N$, $q = 1 - p$,

$$\bar{x} = \frac{1}{N}\sum_{k=1}^{N} x_k, \text{ and } s^2 = \frac{1}{N}\sum_{k=1}^{N} x_k^2.$$

4. *Poisson approximation to binomial:* Let $X_1, X_2, \ldots$ be a sequence of independent binomial random variables, where X_n is binomial$(n, \lambda/n)$ for $n = 1, 2, 3, \ldots$, and $0 < \lambda < 1$. Let Y have a Poisson(λ) distribution. Then, for $x = 0, 1, 2, \ldots$,

$$|\mathbb{P}(X_n \le x) - \mathbb{P}(Y \le x)| \to 0 \text{ as } n \to \infty.$$

References

[Hog78] T. Höglund. Sampling from a finite population: a remainder term estimate. *Scand. J. Stat.*, **5**:69–71, 1978.

Appendix C
Tables

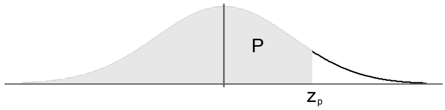

TABLE C.1. Cumulative normal distribution—values of P corresponding to z_p for the standard normal curve.

z_p	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.719	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8707	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

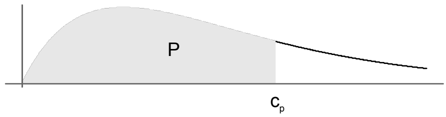

TABLE C.2. Percentiles of the Chisquare distribution—values of c_p corresponding to P.

df	$\chi^2_{.005}$	$\chi^2_{.01}$	$\chi^2_{.025}$	$\chi^2_{.05}$	$\chi^2_{.1}$	$\chi^2_{.9}$	$\chi^2_{.95}$	$\chi^2_{.975}$	$\chi^2_{.99}$	$\chi^2_{.995}$
1	0.000039	0.00016	0.00098	0.0039	0.0158	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.1026	0.2107	4.61	5.99	7.38	9.21	10.60
3	0.0717	0.115	0.216	0.352	0.584	6.25	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	1.064	7.78	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.15	1.61	9.24	11.07	12.83	15.09	16.75
6	0.676	0.872	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55
7	0.989	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.95
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27
18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81	37.16
20	7.43	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57	40.00
24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
120	83.85	86.92	91.57	95.70	100.62	140.23	146.57	152.21	158.95	163.65

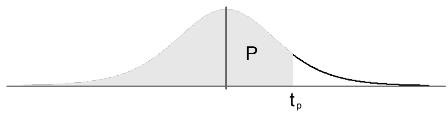

TABLE C.3. Percentiles of the t distribution—values of t_p corresponding to P.

df	$t_{.60}$	$t_{.70}$	$t_{.80}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552	2.878
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528	2.845
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518	2.831
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500	2.807
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492	2.797
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485	2.787
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467	2.763
29	0.256	0.530	0.854	1.311	1.699	2.045	2.462	2.756
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457	2.750
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704
60	0.254	0.527	0.848	1.296	1.671	2.000	2.390	2.660
120	0.254	0.526	0.845	1.289	1.658	1.980	2.358	2.617
∞	0.253	0.524	0.842	1.282	1.645	1.96	2.326	2.576

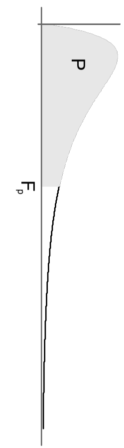

TABLE C.4. Percentiles of the F distribution—values of $F_{.90}$, for k degrees of freedom in the numerator and m degrees of freedom in the denominator.

$m\backslash k$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
1	39.86	49.5	53.59	55.83	57.24	58.2	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.2	5.18	5.18	5.17	5.16	5.15	5.14
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.9	3.87	3.84	3.83	3.82	3.8	3.79	3.78
5	4.06	3.78	3.62	3.52	3.45	3.4	3.37	3.34	3.32	3.3	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.9	2.87	2.84	2.82	2.8	2.78	2.76	2.74
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.7	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.5	2.46	2.42	2.4	2.38	2.36	2.34	2.32
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.3	2.28	2.25	2.23	2.21	2.18
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.2	2.18	2.16	2.13	2.11	2.08
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.3	2.27	2.25	2.21	2.17	2.12	2.1	2.08	2.05	2.03	2.00
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.1	2.06	2.04	2.01	1.99	1.96	1.93
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.2	2.16	2.14	2.1	2.05	2.01	1.98	1.96	1.93	1.9	1.88
14	3.1	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.1	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83
15	3.07	2.7	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.9	1.87	1.85	1.82	1.79
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75
17	3.03	2.64	2.44	2.31	2.22	2.15	2.1	2.06	2.03	2.0	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72
18	3.01	2.62	2.42	2.29	2.2	2.13	2.08	2.04	2.0	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69
19	2.99	2.61	2.4	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.7	1.67
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.0	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.9	1.86	1.81	1.76	1.73	1.7	1.67	1.64	1.6
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.8	1.74	1.72	1.69	1.66	1.62	1.59
24	2.93	2.54	2.33	2.19	2.1	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.7	1.67	1.64	1.61	1.57
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54
27	2.9	2.51	2.3	2.17	2.07	2.0	1.95	1.91	1.87	1.85	1.8	1.75	1.7	1.67	1.64	1.6	1.57	1.53
28	2.89	2.5	2.29	2.16	2.06	2.0	1.94	1.9	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52
29	2.89	2.5	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.5
40	2.84	2.44	2.23	2.09	2.0	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42
60	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.6	1.54	1.51	1.48	1.44	1.4	1.35
120	2.75	2.35	2.13	1.99	1.9	1.82	1.77	1.72	1.68	1.65	1.6	1.55	1.48	1.45	1.41	1.37	1.32	1.26

TABLE C.5. Percentiles of the F distribution—values of $F_{.95}$, for k degrees of freedom in the numerator and m degrees of freedom in the denominator.

$m\backslash k$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
1	161.45	199.5	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.91	245.95	248.01	249.05	250.1	251.14	252.2	253.25
2	18.51	19.0	19.16	19.25	19.3	19.33	19.35	19.37	19.38	19.4	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.7	8.66	8.64	8.62	8.59	8.57	8.55
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.0	5.96	5.91	5.86	5.8	5.77	5.75	5.72	5.69	5.66
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.5	4.46	4.43	4.4
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.1	4.06	4.0	3.94	3.87	3.84	3.81	3.77	3.74	3.7
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.3	3.27
8	5.32	4.46	4.07	3.84	3.69	3.58	3.5	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.9	2.86	2.83	2.79	2.75
10	4.96	4.1	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.7	2.66	2.62	2.58
11	4.84	3.98	3.59	3.36	3.2	3.09	3.01	2.95	2.9	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45
12	4.75	3.89	3.49	3.26	3.11	3.0	2.91	2.85	2.8	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.6	2.53	2.46	2.42	2.38	2.34	2.3	2.25
14	4.6	3.74	3.34	3.11	2.96	2.85	2.76	2.7	2.65	2.6	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18
15	4.54	3.68	3.29	3.06	2.9	2.79	2.71	2.64	2.59	2.54	2.48	2.4	2.33	2.29	2.25	2.2	2.16	2.11
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06
17	4.45	3.59	3.2	2.96	2.81	2.7	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.1	2.06	2.01
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97
19	4.38	3.52	3.13	2.9	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93
20	4.35	3.49	3.1	2.87	2.71	2.6	2.51	2.45	2.39	2.35	2.28	2.2	2.12	2.08	2.04	1.99	1.95	1.9
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.1	2.05	2.01	1.96	1.92	1.87
22	4.3	3.44	3.05	2.82	2.66	2.55	2.46	2.4	2.34	2.3	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84
23	4.28	3.42	3.03	2.8	2.64	2.53	2.44	2.37	2.32	2.27	2.2	2.13	2.05	2.01	1.96	1.91	1.86	1.81
24	4.26	3.4	3.01	2.78	2.62	2.51	2.42	2.36	2.3	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79
25	4.24	3.39	2.99	2.76	2.6	2.49	2.4	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.9	1.85	1.8	1.75
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.2	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73
28	4.2	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71
29	4.18	3.33	2.93	2.7	2.55	2.43	2.35	2.28	2.22	2.18	2.1	2.03	1.94	1.9	1.85	1.81	1.75	1.7
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.0	1.92	1.84	1.79	1.74	1.69	1.64	1.58
60	4.0	3.15	2.76	2.53	2.37	2.25	2.17	2.1	2.04	1.99	1.92	1.84	1.75	1.7	1.65	1.59	1.53	1.47
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.5	1.43	1.35

TABLE C.6. Percentiles of the F distribution—values of $F_{.975}$, for k degrees of freedom in the numerator and m degrees of freedom in the denominator.

$m \backslash k$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
1	647.79	799.5	864.16	899.58	921.85	937.11	948.22	956.66	963.28	968.63	976.71	984.87	993.1	997.25	1001.41	1005.6	1009.8	1014.02
2	38.51	39.0	39.17	39.25	39.3	39.33	39.36	39.37	39.39	39.4	39.41	39.43	39.45	39.46	39.46	39.47	39.48	39.49
3	17.44	16.04	15.44	15.1	14.88	14.73	14.62	14.54	14.47	14.42	14.34	14.25	14.17	14.12	14.08	14.04	13.99	13.95
4	12.22	10.65	9.98	9.6	9.36	9.2	9.07	8.98	8.9	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.36	8.31
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.12	6.07
6	8.81	7.26	6.6	6.23	5.99	5.82	5.7	5.6	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.96	4.9
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.9	4.82	4.76	4.67	4.57	4.47	4.41	4.36	4.31	4.25	4.2
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.3	4.2	4.1	4.0	3.95	3.89	3.84	3.78	3.73
9	7.21	5.71	5.08	4.72	4.48	4.32	4.2	4.1	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.45	3.39
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.2	3.14
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	3.0	2.94
12	6.55	5.1	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.85	2.79
13	6.41	4.97	4.35	4.0	3.77	3.6	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.72	2.66
14	6.3	4.86	4.24	3.89	3.66	3.5	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.61	2.55
15	6.2	4.77	4.15	3.8	3.58	3.41	3.29	3.2	3.12	3.06	2.96	2.86	2.76	2.7	2.64	2.59	2.52	2.46
16	6.12	4.69	4.08	3.73	3.5	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.45	2.38
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.5	2.44	2.38	2.32
18	5.98	4.56	3.95	3.61	3.38	3.22	3.1	3.01	2.93	2.87	2.77	2.67	2.56	2.5	2.44	2.38	2.32	2.26
19	5.92	4.51	3.9	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.27	2.2
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.22	2.16
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.8	2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.18	2.11
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.7	2.6	2.5	2.39	2.33	2.27	2.21	2.14	2.08
23	5.75	4.35	3.75	3.41	3.18	3.02	2.9	2.81	2.73	2.67	2.57	2.47	2.36	2.3	2.24	2.18	2.11	2.04
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.7	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.08	2.01
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.3	2.24	2.18	2.12	2.05	1.98
26	5.66	4.27	3.67	3.33	3.1	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.22	2.16	2.09	2.03	1.95
27	5.63	4.24	3.65	3.31	3.08	2.92	2.8	2.71	2.63	2.57	2.47	2.36	2.25	2.19	2.13	2.07	2.0	1.93
28	5.61	4.22	3.63	3.29	3.06	2.9	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.98	1.91
29	5.59	4.2	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.96	1.89
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.2	2.14	2.07	2.01	1.94	1.87
40	5.42	4.05	3.46	3.13	2.9	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.8	1.72
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.67	1.58
120	5.15	3.8	3.23	2.89	2.67	2.52	2.39	2.3	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.53	1.43

TABLE C.7. Percentiles of the F distribution—values of $F_{.99}$, for k degrees of freedom in the numerator and m degrees of freedom in the denominator.

$m\backslash k$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120
1	4052.18	4999.5	5403.35	5624.58	5763.65	5858.99	5928.36	5981.07	6022.47	6055.85	6106.32	6157.28	6208.73	6234.63	6260.65	6286.78	6313.03	6339.39
2	98.5	99.0	99.17	99.25	99.3	99.33	99.36	99.37	99.39	99.4	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.6	26.5	26.41	26.32	26.22
4	21.2	18.0	16.69	15.98	15.52	15.21	14.98	14.8	14.66	14.55	14.37	14.2	14.02	13.93	13.84	13.75	13.65	13.56
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.2	9.11
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.1	7.98	7.87	7.72	7.56	7.4	7.31	7.23	7.14	7.06	6.97
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.2	5.12	5.03	4.95
9	10.56	8.02	6.99	6.42	6.06	5.8	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.4
10	10.04	7.56	6.55	5.99	5.64	5.39	5.2	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.0
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.4	4.25	4.1	4.02	3.94	3.86	3.78	3.69
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.5	4.39	4.3	4.16	4.01	3.86	3.78	3.7	3.62	3.54	3.45
13	9.07	6.7	5.74	5.21	4.86	4.62	4.44	4.3	4.19	4.1	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.8	3.66	3.51	3.43	3.35	3.27	3.18	3.09
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.0	3.89	3.8	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96
16	8.53	6.23	5.29	4.77	4.44	4.2	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.1	3.02	2.93	2.84
17	8.4	6.11	5.18	4.67	4.34	4.1	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.0	2.92	2.83	2.75
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.6	3.51	3.37	3.23	3.08	3.0	2.92	2.84	2.75	2.66
19	8.18	5.93	5.01	4.5	4.17	3.94	3.77	3.63	3.52	3.43	3.3	3.15	3.0	2.92	2.84	2.76	2.67	2.58
20	8.1	5.85	4.94	4.43	4.1	3.87	3.7	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.4	3.31	3.17	3.03	2.88	2.8	2.72	2.64	2.55	2.46
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.5	2.4
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.3	3.21	3.07	2.93	2.78	2.7	2.62	2.54	2.45	2.35
24	7.82	5.61	4.72	4.22	3.9	3.67	3.5	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.4	2.31
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.7	2.62	2.54	2.45	2.36	2.27
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.5	2.42	2.33	2.23
27	7.68	5.49	4.6	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.2
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.9	2.75	2.6	2.52	2.44	2.35	2.26	2.17
29	7.6	5.42	4.54	4.04	3.73	3.5	3.33	3.2	3.09	3.0	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14
30	7.56	5.39	4.51	4.02	3.7	3.47	3.3	3.17	3.07	2.98	2.84	2.7	2.55	2.47	2.39	2.3	2.21	2.11
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.8	2.66	2.52	2.37	2.29	2.2	2.11	2.02	1.92
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.5	2.35	2.2	2.12	2.03	1.94	1.84	1.73
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53

Index

Nguyen and Rogers: Fundamentals of Mathematical Statistics: Volume I: Probability for Statistics
Nguyen and Rogers: Fundamentals of Mathematical Statistics: Volume II: Statistical Inference
Noether: Introduction to Statistics: The Nonparametric Way
Nolan and Speed: Stat Labs: Mathematical Statistics Through Applications
Peters: Counting for Something: Statistical Principles and Personalities
Pfeiffer: Probability for Applications
Pitman: Probability
Rawlings, Pantula and Dickey: Applied Regression Analysis
Robert: The Bayesian Choice: A Decision-Theoretic Motivation
Robert and Casella: Monte Carlo Statistical Methods
Santner and Duffy: The Statistical Analysis of Discrete Data
Saville and Wood: Statistical Methods: The Geometric Approach
Sen and Srivastava: Regression Analysis: Theory, Methods, and Applications
Shao: Mathematical Statistics
Shumway and Stoffer: Time Series Analysis and Its Applications
Terrell: Mathematical Statistics: A Unified Introduction
Whittle: Probability via Expectation, Fourth Edition
Zacks: Introduction to Reliability Analysis: Probability Models and Statistical Methods