Review of Statistics [1]

Jasmine(Yu) Hao

Faculty of Business and Economics
Hong Kong University

September 2, 2021

[1] This section is based on Stock and Watson (2020), Chapter 3.

Estimators

Suppose you want to understand the distribution of X in the population.

- When a statistic $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$ is a function of an i.i.d. sample, then the distribution is determined by the population distribution is F and the sample size is n.
- We call the distribution of $\hat{\theta}$ the **sample distribution**.

The goal of an estimator $\hat{\theta}$ is to learn about the parameter θ, we evaluate the

- The exact bias and variance.
- The distribution under normality.
- The asymptotic distribution as $n \to \infty$.

Goodness of Estimators

Let $\hat{\theta}$ be an estimator of θ. Then

- The bias of $bias(\hat{\theta})$ is $E[\hat{\theta}] - \theta$.
  - We say an estimator is **unbiased** if the bias is 0.
- The **mean squared error** of an estimator $\hat{\theta}$ for θ is

$$mse(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)^2].$$

  - The mean squared error is $mse(\hat{\theta}) = \mathrm{var}(\hat{\theta}) + (bias(\hat{\theta}))^2$.

Best Unbiased Estimator

Definition 1 (Best Linear Unbiased Estimator (BLUE))

If $\sigma^2 < \infty$ the sample mean $\bar{X}_n$ has the lowest variance among all linear unbiased estimators of μ.

Bias, consistency, and efficiency

- Suppose $Y_1, \ldots, Y_n$ are i.i.d.
- Denote an estimator for μ_Y as $\hat{\mu}_Y$.
- The bias of $\hat{\mu}_Y$ is $E(\hat{\mu}_Y) - \mu_Y$.
- $\hat{\mu}_Y$ is an **unbiased** estimator of $\hat{\mu}_Y$ if $E(\hat{\mu}_Y) = \mu_Y$.
- $\hat{\mu}_Y$ is an **consistent** estimator of $\hat{\mu}_Y$ if $\hat{\mu}_Y \rightarrow_p \mu_Y$.
- Let $\tilde{\mu}_Y$ denote another estimator for μ_Y, and suppose both $\bar{\mu}_Y$ and $\tilde{\mu}_Y$ are consistent. Then $\hat{\mu}_Y$ is more efficient if $var(\hat{\mu}_Y) < var(\tilde{\mu}_Y)$.

Properties

- $E(\bar{Y}) = \mu_Y$, so $\bar{Y}$ is an unbiased estimator of μ_Y.
- the law of large numbers states that $\bar{Y} \to_p \mu_Y$, $\bar{Y}$ is consistent.
- Consider $\tilde{Y} = \frac{1}{n}(\frac{1}{2}Y_1 + \frac{3}{2}Y_2 + \dots)$, then $var(\tilde{Y}) = 1.25\sigma_Y^2/n > var(\bar{Y}) = \sigma_Y^2/n$. $\bar{Y}$ is more efficient than $\tilde{Y}$.
- $\bar{Y}$ is the Best Linear Unbiased Estimator for μ_Y.

Hypothesis

- A point hypothesis is the statement that θ equals a specific value θ_0.
- A common example is θ measures the effect the proposed policy. A typical question is whether $\theta = \theta_0$.
- The **null hypothesis**, written as $H_0 : \theta = \theta_0$.
- The **alternative hypothesis**, written as $H_A : \theta \neq \theta_0$, is the set $\{\theta \in \Theta : \theta \neq \theta_0\}$.
  - **One-sided** hypothesis: $H_A : \theta > \theta_0$.
  - **Two-sided** hypothesis: $H_A : \theta \neq \theta_0$.

Acceptance and Rejection

- A hypothesis test is a decision based on data. We can either **fail to reject** the null hypothesis or **reject** the null hypothesis.
- An alternative way to express a decision rule is to construct a real-valued function of the data called a **test statistics**

$$T = T(X_1, \ldots, X_n)$$

  together with a **critical region** C.
- A hypothesis can be expressed as
  - Fail to reject H_0 if $T \in C$.
  - Reject H_0 if $T \notin C$.

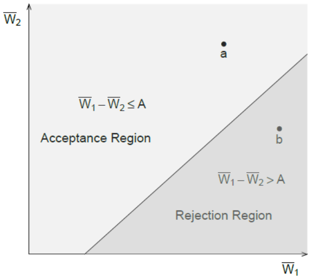
Note: "Accept" H_0 does not mean H_0 is true.

Example - Hypothesis Testing I

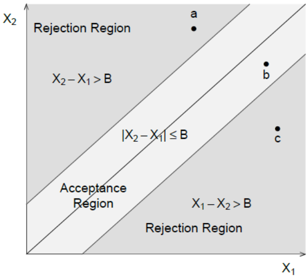
Consider the following examples:

- $2n$ adults who were raised in similar settings, n attended early childhood education. Let $\bar{W}_1$ be the average wage in the early childhood education group, and let $\bar{W}_2$ be the average wage in the remaining sample. Null hypothesis $H_0 : \bar{W}_1 > \bar{W}_2$.
- You ride each bus once and record the time it takes to travel from home to the university. Let X_1 and X_2 be the two recorded travel times. You adopt the following decision rule: If the absolute difference in travel times is greater than B minutes you will reject the hypothesis that the average travel times are the same, otherwise you will accept the hypothesis.

Example - Hypothesis Testing II



(a) Early Childhood Education Example



(b) Bus Travel Example

Type I and Type II error

- A false rejection of the null hypothesis is a **Type I error**.
- A false acceptance of the alternative hypothesis is a **Type II error**.

	Accept H_0	Reject H_0
H_0 true	Correct Decision	Type I Error
H_1 true	Type II Error	Correct Decision

P-value I

- the sample average $\bar{Y}$ will rarely be exactly equal to the hypothesized value $\mu_{Y,0}$.
- Differences between $\bar{Y}$ and $\mu_{Y,0}$ can arise because
  - the true mean is not $\mu_{Y,0}$ (the null hypothesis is false) or
  - the true mean equals $\mu_{Y,0}$ (the null hypothesis is true) but $\bar{Y}$ differs from $\mu_{Y,0}$ because of random sampling.
- impossible to distinguish between these two possibilities with certainty.

P-value II

With a sample of data

- cannot conclude if H_0 Is true.
- can do **probabilistic calculation** that permits testing the null hypothesis in a way that accounts for sampling uncertainty.
- How?compute the p-value of the null hypothesis.

P-value III

- The p-value, also called the **significance probability**, is the probability of drawing a statistic **at least as adverse to the null hypothesis** as the one you actually computed in your sample, assuming the null hypothesis is correct.
- In the case at hand, the p-value is the probability of drawing $\bar{Y}$ at least as far in the tails of its distribution under the null hypothesis as the sample average you actually computed.
- $p - \text{value} = Pr(|\bar{Y} - \mu_{Y,0}| > |\bar{Y}^{act} - \mu_{Y,0}|)$

Example I

In a sample of recent college graduates, the average wage is \$22.64. The p-value is the probability of observing a value of Y at least as different from \$20 (the population mean under the null hypothesis) as the observed value of \$22.64 by pure random sampling variation, assuming that the null hypothesis is true.

- If this p-value is small (say, 0.1%), unlikely that this sample drawn if the null hypothesis is true;
  - reasonable to conclude that the null hypothesis is not true.
- if this p-value is large (say, 40%), likely that the observed sample average of \$22.64 could have arisen just by random sampling variation if the null hypothesis is true;
  - the evidence against the null hypothesis is weak in this probabilistic sense,(**fail to reject**)

Example II

The p-value is the area in the tails of the distribution of $\bar{Y}$ under the null hypothesis beyond $\mu_{Y,0} \pm |\bar{Y}^{act} - \mu_{Y,0}|$.

- To compute the p-value, need to know the shape of the distribution.
- With CLT, the sampling distribution of $\bar{Y}$ is well approximated by a normal distribution

When σ_Y is known, then we can compute the *p*-value

- Recall: By CLT, $(\bar{Y} - \mu_Y)/\sqrt{\sigma_{\bar{Y}}} \rightarrow_d N(0,1)$, then $\sqrt{n}(\bar{Y} - \mu_Y) \rightarrow_d N(0, \sigma_Y^2)$.
- Under the null hypothesis, $p - \text{value} = Pr(\left|\frac{\bar{Y} - \mu_{Y,0}}{\sigma_{\bar{Y}}}\right| > \left|\frac{\bar{Y}^{act} - \mu_{Y,0}}{\sigma_{\bar{Y}}}\right|) = 2\Phi(-\left|\frac{\bar{Y}^{act} - \mu_{Y,0}}{\sigma_{\bar{Y}}}\right|)$,
- where Φ is the standard normal cumulative distribution function.

Computing p-value

- Suppose we are interested in testing the null hypothesis in $H_0 : \mathbb{E}(X) = \mu$ with the alternative hypothesis $H_A : \mathbb{E}(X) \neq \mu$.
  - Two-sided test.
- We observe the realization of $X_1, \ldots, X_n$ as $x_1, \ldots, x_n$.
- Note that $\bar{X}$ is a function of $X_1, \ldots, X_n$, which are i.i.d., therefore is a random variable.
  - Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
  - and $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.
- Under H_0, the distribution of $\frac{\bar{X}-\mathbb{E}(X)}{\sigma_{\bar{X}}} \sim N(0,1)$(CLT).
- $p = 1 - \mathbb{P}\left(\left|\frac{\bar{X}-\mathbb{E}(X)}{\sigma_{\bar{X}}}\right| < \left|\frac{\bar{x}-\mathbb{E}(X)}{\sigma_{\bar{X}}}\right|\right)$.

Issue: $\sigma_{\bar{X}}$ **unknown.**

Sample Variance

If the following assumptions hold:

1. $X_1, \ldots, X_n$ are i.i.d.
2. $\mathbb{E}(X_i) < \infty$.

The sample variance is computed

$$\bar{s}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

- μ is unknown, need to be estimated.
- $\mathbb{E}((X - \bar{X})^2) \to \frac{n-1}{n}\sigma$.
- The sample variance is a consistent estimator of the population variance.

Sample Variance - Example

- When $Y_1, \ldots, Y_n$ are i.i.d. draws from a Bernoulli distribution with success probability p,
- the formula for the variance of $\bar{Y}$ simplifies to $p(1-p)/n$,
- The formula for the standard error also takes on a simple form that depends only on Y and n: $SE(\bar{Y}) = \sqrt{\bar{Y}(1-\bar{Y})} > n$.

t-statistic

The standardized sample average can be constructed using

$$t = \frac{\bar{X} - \mu}{\sqrt{\bar{s}^2}}.$$

With the sample of $x_1, \ldots, x_n$, we can compute the sample t-statistic t^{sample}.

The p-value is given by

$$p\text{-value} = 2\Phi(-|t^{sample}|).$$

Significance Level

When construct hypothesis test, can fix a significance level.

- ▷ α-significance test means the tolerance to make Type I error is α.
- ▷ α is referred to as the **size** of the test.

Suppose the two-sided test has the **significance level** of α, the rule is "**Reject** H_0 **if** $|t^{sample}| > 1 - \Phi^{-1}(\alpha/2)$".

- ▷ $\alpha = 1\%$, $1 - \Phi^{-1}(\alpha/2) = 2.58$.
- ▷ $\alpha = 5\%$, $1 - \Phi^{-1}(\alpha/2) = 1.96$.
- ▷ $\alpha = 10\%$, $1 - \Phi^{-1}(\alpha/2) = 1.64$.

Confidence Interval I

- random sampling error makes it impossible to learn the exact value of the population mean
- it is possible to **construct a set** of values that contains the true population mean with a certain prespecified probability.
- It's called **confidence set**, the prespecified probability that μ_Y is contained in this set is called the **confidence level**.
- The confidence set for μ_Y turns out to be all the possible values of the mean between a lower and an upper limit, so that the confidence set is an interval, called a **confidence interval**.
- The **coverage probability** of a confidence interval for the population mean is the probability, computed over all possible random samples, that it contains the true population mean.

Confidence Interval II

A 95\% two-sided confidence interval for mY is an interval constructed so that it contains the true value of mY in 95% of all possible random samples. When the sample size n is large, 90%, 95%, and 99% confidence intervals for mY are:

- 90% confidence interval for $\mu_Y = \bar{Y} \pm 1.64 SE(\bar{Y})$
- 95% confidence interval for $\mu_Y = \bar{Y} \pm 1.96 SE(\bar{Y})$
- 99% confidence interval for $\mu_Y = \bar{Y} \pm 2.58 SE(\bar{Y})$

Example

- consider the problem of constructing a 95% confidence interval for the mean hourly earnings of recent college graduates using a hypothetical random sample of 200 recent college graduates where
- $\bar{Y} = \$22.64$ and $se(\bar{Y}) = 1.28$.
- The 95% confidence interval for mean hourly earnings is $22.64 \pm 1.96 \times 1.28 = (20.13, 25.15)$.
- The **coverage probability** of a confidence interval for the population mean is the probability, computed over all possible random samples, that it contains the true population mean

Confidence Interval(More general sense) I

We are interested in learning a parameter of interest θ from i.i.d. random sample of $X_1, \ldots, X_n$.

- With random sampling error, it's impossible to learn the exact value of the parameter of interest.
- Construct a **confidence set**: the parameter of interested has $1 - \alpha$ probability to fall into the confidence set.
- The **coverage probability** of the interval estimator is the probability that the random interval contains the true parameter.
  - An $1 - \alpha$ **asymptotic confidence interval** for a parameter has the **asymptotic coverage probability** $1 - \alpha$.

Confidence Interval(More general sense) II

A normal-based $1-\alpha$ confidence interval is

$$CI = [\hat{\theta} - Z_{1-\alpha/2} s(\hat{\theta}), \hat{\theta} + Z_{1-\alpha/2} s(\hat{\theta})],$$

where $\hat{\theta}$ is the estimator for θ and $se(\hat{\theta})$ is the estimated standard deviation. $Z_{1-\alpha/2}$ is the $1-\alpha/2$-quantile of a normal distribution.

Test for Difference Between Two Groups I

Suppose we observe the i.i.d sample $W_1, \ldots, W_{n_1}, \ldots, W_n$.

- Sample $W_1, \ldots, W_{n_1}$ are the monthly wage of graduates with master's degree, let μ_1 denote the population mean and σ_1^2 the population variance of group 1.
- Sample $W_{n_1+1}, \ldots, W_n$ are the monthly wage of graduates with bachelor's degree, let μ_2 denote the population mean and σ_2^2 the population variance of group 2.
- Let $n_2 = n - n_1$.
- $H_0 : \mu_1 - \mu_2 > d_0$, $H_1 : \mu_1 - \mu_2 \leq d_0$, with significance level of α.

Test for Difference Between Two Groups II

- The parameter of interest is $\theta = \mu_1 - \mu_2$.
- Let $\bar{W}_1$ and $\bar{W}_2$ be the estimated sample mean and s_1^2 and s_2^2 be the estimated sample variance for group 1 and group 2.
- The standard error of $\hat{\theta} = \bar{W}_1 - \bar{W}_2$ is $se(\hat{\theta}) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$.
- We construct the t-statistic as $t = \frac{\hat{\theta} - d_0}{se(\hat{\theta})}$.
- We reject H_0 if $t > Z_{1-\alpha}$.

Social Class or Education? Childhood Circumstances and Adult Earning I

This example is based on the example in SW2020 Chapter 3, p.p. 122.

TABLE 3.1 Differences in Household Income According to Childhood Socioeconomic Circumstances, Grouped by Level of Highest Qualification

	Father's NS-SEC = Higher			Father's NS-SEC = Routine			Difference, Higher vs. Routine			
Qualification	$\bar{Y}_h$	s_h	n_h	$\bar{Y}_r$	s_r	n_r	$\bar{Y}_h - \bar{Y}_r$	$SE(\bar{Y}_h - \bar{Y}_r)$	95% Confidence Interval for d	
None	£2,223.13	£2,115.12	1129	£1,842.98	£1,48729	6383	£380.15	£65.64	£251.38	£508.93
GCSE/O-Level	£2,83718	£1,819.73	1962	£2,596.93	£1,738.47	4042	£240.25	£49.35	£143.49	£33700
A-Level	£3,045.99	£2,451.81	1216	£2,745.70	£1,912.50	1169	£300.30	£89.85	£124.11	£476.49
Undergraduate degree or more	£3,690.51	£2,743.55	4359	£3,370.96	£2,443.58	2505	£319.55	£64.11	£193.86	£445.23
All categories	£3,215.71	£2,49773	8666	£2405.45	£1,886.86	14099	£810.25	£31.18	£749.13	£871.38

Source: Understanding Society.

Figure

Social Class or Education? Childhood Circumstances and Adult Earning II

- breaks down the differences in mean household income for individuals according to their are father's NS-SEC occupation type,
- and considers these differences for selected highest level of educational qualification
- The data shows that within both groups according to the NS-SEC of a father's occupation, those with higher qualifications are part of households with higher total income.
- Test the differences between mean income by the father's occupational categorization $(Y_h - Y_r)$ for each of the educational groupings.

Social Class or Education? Childhood Circumstances and Adult Earning III

- For individuals with no qualifications
- test statistics $= \frac{(2223.13-1842.98)}{2115.12^2/1129+1487.29^2/6383} = 5.7911$.
- The 95 \% CI for the difference $(Y_h - Y_r)$ is $(2223.13 - 1842.98) \pm 1.96\sqrt{2115.12^2/1129 + 1487.29^2/6383} = (251.38, 508.93)$.

Use t-distribution when n is small

Consider the t-statistic used to test the hypothesis $H_0 : \mu_Y = \mu_{Y,0}$, using data $Y_1, \ldots, Y_n$.

$$t = \frac{\bar{Y} - \mu_{Y,0}}{\sqrt{s_Y^2/n}},$$

where $s_Y^2$is the estimated sample mean.

- When n is larger, under general conditions the t-statistic has a standard normal distribution if the sample size is large and the null hypothesis is true.
- When n is small, then the t-statistic in Equation has a Student t distribution with n - 1 degrees of freedom.

Example I

To illustrate a test for the difference between two means

- let μ_w be the mean hourly earnings in the population of women recently graduated from college,
- let μ_m be the population mean for recently graduated men.
- Consider the null hypothesis that mean earnings for these two populations differ by a certain amount, say, d_0. Then the null hypothesis and the two-sided alternative hypothesis are $H_0 : \mu_m - \mu_w = d_0$ vs. $H_1 : \mu_m - \mu_w \neq d_0$

Example II

Population means are unknown:

- must be estimated from samples of men and women.
- Suppose we have samples of n_m men and n_w women drawn at random from their populations.
- Let the sample average annual earnings be $\bar{Y}_m$ for men and $\bar{Y}_w$ for women.
- An estimator of $\mu_m - \mu_w$ is $\bar{Y}_m - \bar{Y}_w$.

Example III

In asymptotics:

- $\bar{Y}_m \to_d N(\mu_m, \sigma_m^2/n_m)$, $\bar{Y}_w \to_d N(\mu_, \sigma_w^2/n_w)$
- By properties of random distributions, $\bar{Y}_m - \bar{Y}_w \to N(\mu_m - \mu_w, (\sigma_m^2/n_m) + (\sigma_w^2/n_w)^2)$.

Example IV

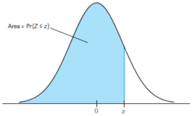
Construct t-statistics:
$t = \frac{\bar{Y}_m - \bar{Y}_w}{\sqrt{(\sigma_m^2/n_m) + (\sigma_w^2/n_w)^2}}$.

- When n is larger(> 30), $t \to_d N(0,1)$.
  - If the test is at 5% significance level, reject if $t > 1.64$.
- When n is small(≤ 30), $t \sim t - \text{dist}_{n-1}$

Distribution Table I

Appendix

TABLE 1 The Cumulative Standard Normal Distribution Function, $\Phi(z) = \Pr(Z \le z)$



	Second Decimal Value of z									
z	0	1	2	3	4	5	6	7	8	9
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611

Distribution Table II

z	0	1	2	3	4	5	6	7	8	9
	Second Decimal Value of z									
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
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

This table can be used to calculate $\Pr(Z \leq z)$ where Z is a standard normal variable. For example, when $z = 1.17$, this probability is 0.8790, which is the table entry for the row labeled 1.1 and the column labeled 7.

References I

Stock, J. H. and Watson, M. W. (2020). *Introduction to econometrics*, volume 4. Pearson New York.