COURSE SUMMARY NOTES

METHODS UNIT 4

Contents

1 Logarithmic Functions

Definition

If $b^x = n$ then

$$\log_b n = x \tag{1.1}$$

Logarithm Laws

$$\log_b (nm) = \log_b (n) + \log_b (m) \tag{1.2}$$

$$\log_b \left(\frac{n}{m}\right) = \log_b (n) - \log_b (m) \tag{1.3}$$

$$\log_b \left(n^k\right) = k \log_b (n) \tag{1.4}$$

$$\log_b (b) = 1 \tag{1.5}$$

$$\log_b (1) = 0 \tag{1.6}$$

$$\log_b \left(\frac{1}{n}\right) = -\log_b (n) \tag{1.7}$$

Natural Logarithm

$$\log_{\mathrm{e}} (n) = \ln (n) \tag{1.8}$$

Changing the Base of Logarithms

$$\log_b n = \frac{\log_c n}{\log_c b} \tag{1.9}$$

Note:

1. Prove the following result by taking $\log_c$ of both sides of $b^x = n$

Solving Equations using Logarithms

When solving for an unknown in the exponent you can take the $\log_b$ of both sides of the equation.

Note:

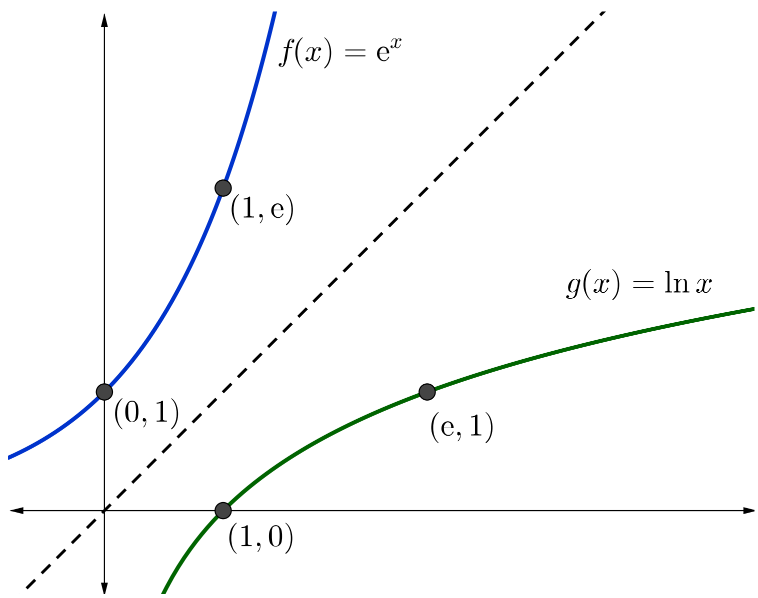
1. It makes no difference what base logarithm you chose to use however some bases may make the result more concise.

$$n^x = m$$

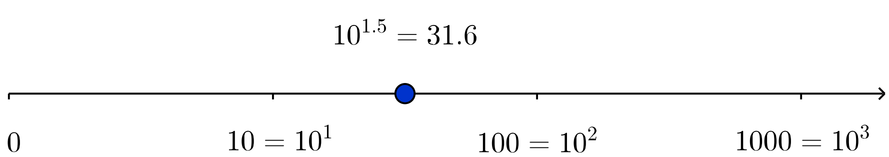
$$\ln(n^x) = \ln(m)$$

$$x = \frac{\ln(m)}{\ln(n)}$$

Logarithm Graph



Logarithmic Scales



2 Calculus Involving Logarithmic Functions

Derivative of $\ln x$

$$\frac{\mathrm{d}}{\mathrm{d}x}(\ln x) = \frac{1}{x} \tag{2.1}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}(\ln f(x)) = \frac{f'(x)}{f(x)} \tag{2.2}$$

Note:

1. It is worth taking note of the domain of $\ln x$ being $x > 0$ as that domain also applies to the derivative.

Extracurricular:

1. Proving 2.1 is an interesting exercise in limits and is useful in the proof of the derivative of e^x.

Integral of $\ln x$

$$\int \frac{1}{x}\,\mathrm{d}x = \ln|x| + c \tag{2.3}$$

$$\int \frac{f'(x)}{f(x)}\,\mathrm{d}x = \ln|f(x)| + c \tag{2.4}$$

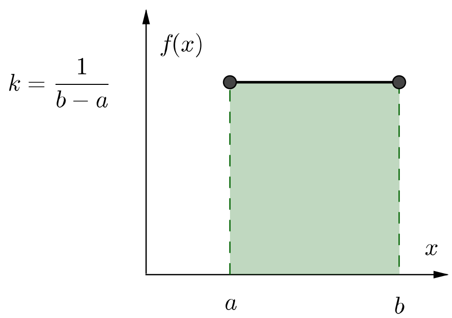
3 Continuous Random Variables

Continuous Uniform Distribution

$$X \sim \mathrm{U}(a, b) \tag{3.1}$$

Note:

1. It can be helpful to view many questions related to Uniform Distribution as simply finding the area of rectangles.

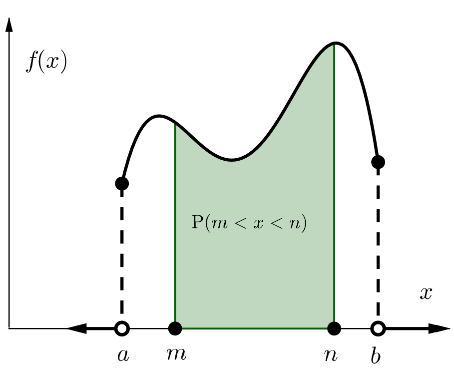


Non-Uniform Distribution

Given

$$P(X = x) = \begin{cases} f(x) & \text{for } a \leq x \leq b \\ 0 & \text{for all other values of } x \end{cases}$$

$$P(m < x < n) = \int_m^n f(x)\, dx \tag{3.2}$$



Expected Value

$$\mathrm{E}(X) = \int_a^b x \cdot f(x)\,\mathrm{d}x \tag{3.3}$$

Note:

1. a is the lower boundary and b is the upper boundary of x for the PDF.

Variance

$$\mathrm{Var}(X) = \int_a^b f(x) \cdot (x-\mu)^2\,\mathrm{d}x \tag{3.4}$$

$$= \int_a^b f(x) \cdot x^2\,\mathrm{d}x - \mu^2 \tag{3.5}$$

$$= E(X^2) - [E(X)]^2 \tag{3.6}$$

Cumulative Distribution Function

$$P(X \leq x) = F(x) = \begin{cases} 0 & x < a \\ \int_a^x f(t)\,\mathrm{d}t & a \leq x \leq b \\ 1 & b < x \end{cases}$$

4 The Normal Distribution

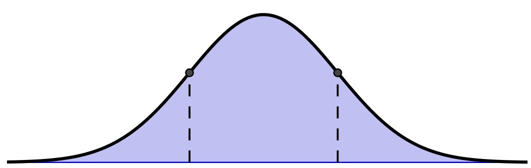
Normal Distribution Function

$$X \sim \mathrm{N}\left(\mu, \sigma^2\right) \tag{4.1}$$

$$f(x) = \frac{1}{\sqrt{2\sigma^2\pi}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4.2}$$

Note:

1. One standard deviation above and below the mean occurs at the points of inflection.



$$\mathrm{P}(\mu - \sigma \le X \le \mu + \sigma) \approx 0.68 \tag{4.3}$$

$$\mathrm{P}(\mu - 2\sigma \le X \le \mu + 2\sigma) \approx 0.95 \tag{4.4}$$

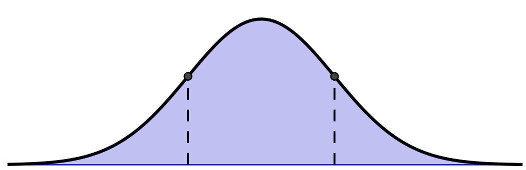
$$\mathrm{P}(\mu - 3\sigma \le X \le \mu + 3\sigma) \approx 0.997 \tag{4.5}$$

Standard Normal Distribution

$$Z \sim N(0, 1) \tag{4.6}$$

Note:

1. The Z score can be interpreted as the number of standard deviations above or below the mean.



Standard Scores

$$z = \frac{x - \mu}{\sigma} \tag{4.7}$$

Quantiles and Percentiles

Given

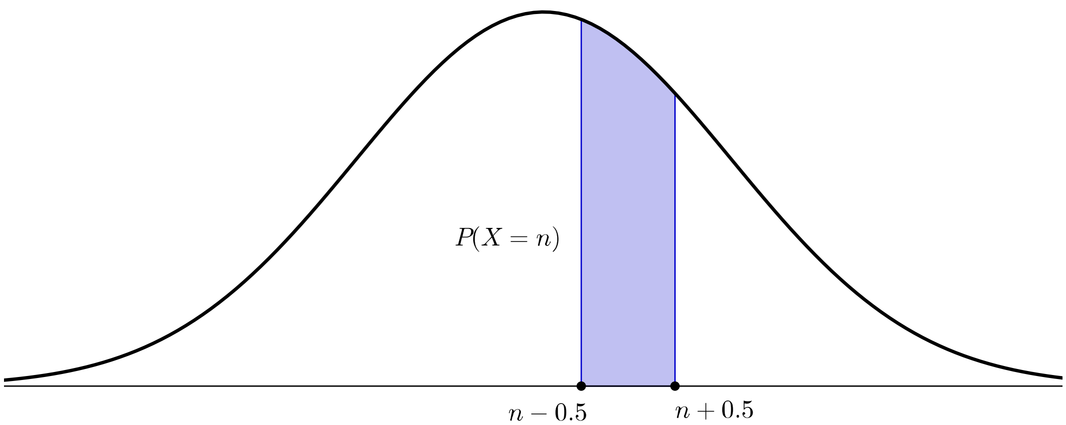
$$P(X < q) = p$$

q is the p **quantile**.
q is the $p \times 100\%$ **percentile**.

Correction for Continuity

We can use a continuous random distribution like the normal distribution to approximate discrete random variables if we make a correction for continuity.

$$P(X = n) \approx P(n - 0.5 \leq X \leq n + 0.5) \tag{4.8}$$



5 Random Sampling

Terms

Sample: part of the population.

Census: data collected from the whole population.

Parameter: characteristic of population, such as mean.

Statistic: estimate of a parameter obtained using a sample.

Survey: obtains the same information from each member of the sample or population.

Fair Sample: representative of the population.

Biased Sample: not representative of the population.

Random Sample: ensures each member of the population has equal chance of being chosen.

Bias

Selection Bias: arises from the choice of the sample. Solution: random sampling.

Design Flaw Bias: arises from faults in the design of the survey. Solution: objective questions.

Interviewer Bias: arises from interviewer technique. Solution: standard questions and question options.

Recall Bias: arises from inaccurate recollection of past events.

Reporting Bias: arises from the selective revealing or suppression of information.

Completion Bias: occurs from surveys being incomplete.

Non-Response Bias: occurs when some subjects do not take part in the survey. The most extreme example of non-response bias is self-selected bias.

Types of Sampling

Systematic or Interval Sampling: choosing representatives from every n^{th} item from the population.

Stratified Sampling: choosing representatives from identified groups of the population in proportion to the size of each group.

Cluster Sampling: choosing groups from the population at random.

Self-selection Sampling: members of the population volunteering themselves rather than being selected.

Capture-Recapture

Let

$$\begin{aligned} N &= \text{size of population} \\ n_1 &= \text{size of first sample} \\ n_2 &= \text{size of second sample} \\ k &= \text{number of tagged within second sample} \end{aligned}$$

$$\frac{n_1}{N} = \frac{k}{n_2} \tag{5.1}$$

6 *Sample Proportions*

Terms

x: the number of a elements with a particular characteristic in the sample.

p: the proportion of a particular section of the population.

$\hat{p}$: the proportion of the particular section of a sample from the population.

n: sample size.

k: a chosen number of standard deviations above and below the mean to achieve a required level of confidence.

Distribution of x

$$x \sim Bin\,(n, p) \tag{6.1}$$

Distribution of $\hat{p}$

As $n \to \infty$

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right) \tag{6.2}$$

Note:

1. $E(\hat{p}) = p$
2. $s_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$
3. $np > 5$ and $n(1-p) > 5$

Extracurricular:

1. This concept is an application of The Central Limit Theorem.

Confidence Intervals

$$p - k\sqrt{\frac{p(1-p)}{n}} \leq \hat{p} \leq p + k\sqrt{\frac{p(1-p)}{n}} \tag{6.3}$$

$$\hat{p} - k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \tag{6.4}$$

Choosing Sample Size

$$E = k\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \tag{6.5}$$

Note:

1. If proportion is not given assume $p = 0.5$.
2. Always round up for n.

Confidence Levels

$$90\% \text{ confidence level } k = 1.645 \tag{6.6}$$

$$95\% \text{ confidence level } k = 1.960 \tag{6.7}$$

$$99\% \text{ confidence level } k = 2.576 \tag{6.8}$$