- Mean (or average)

$$\mu = \frac{\sum_{i=1}^n x_i}{n}$$

- Weighted mean

$$\text{weighted mean}: \quad \bar{x} = \frac{\sum_{i=1}^n w_i \cdot x_i}{\sum_{i=1}^n w_i}$$

- The **median** of a data set is the measure of center that is the *middle value* when the original data values are arranged in order of increasing magnitude.
- The **mode** of a data set is the value that occurs with the greatest frequency.
- Sample variance

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

- Sample standard deviation

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

- z-score

$$z = \frac{x - \bar{x}}{s}$$

- Classical probability

$$P(A) = \frac{\text{number of ways } A \text{ occurred}}{\text{number of different simple events}}$$

- Disjoint: Two events are disjoint if

$$P(A \cap B) = 0$$

- Inclusion-Exclusion

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

- Complement

$$P(A^c) = 1 - P(A)$$

- Independence: Two events are independent if

$$P(AB) = P(A)P(B)$$

- De Morgan's Law

$$(A \cup B)^c = A^c \cap B^c$$

- Probability of at least one event

$$P(k \geq 1) = 1 - P(k = 0)$$

- Conditional probability

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

- Total probability: If we observe event B after event A, then

$$P(B) = P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)$$

- Bayes' Rule

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)}$$

- Expected value

$$\mu = \mathrm{E}[X] = \sum_{i=1}^n x_i \cdot P(x_i)$$

- Second moment

$$\mathrm{E}[X^2] = \sum_{i=1}^n x_i^2 \cdot P(x_i)$$

- Variance

$$\sigma^2 = \mathrm{Var}[X] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$$

- Range rule of thumb

$$(\mu - 2\sigma, \mu + 2\sigma)$$

- Factorial

$$n! = \prod_{i=1}^n i$$

- Choose

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

- Binomial distribution mass function

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

  1. $\mu = np$
  2. $\sigma^2 = np(1-p)$

- Poisson distribution mass function

$$P(k) = \frac{\mu^k \cdot e^{-\mu}}{k!}$$

1. $\mathrm{E}[X] = \mu$
2. $\mathrm{Var}[X] = \mu$

- Geometric distribution mass function (with k "failures" until "success" on the $(k+1)^{\text{st}}$ iteration)

$$P(k) = (1-p)^k p$$

1. $\mathrm{E}[X] = \frac{1-p}{p}$
2. $\mathrm{Var}[X] = \frac{1-p}{p^2}$

- Memoryless property

$$P(X > m+n | X > n) = P(X > m)$$

- arrangements (selecting r items out of n)

	with replacement	without replacement
combinations	$\left(\binom{n}{r}\right) = \frac{(n+r-1)!}{r!(n-1)!}$	$\binom{n}{r} = \frac{n!}{r!(n-r)!}$
permutations	n^r	$\frac{n!}{(n-r)!}$

- Correlation

$$r = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{(n-1)s_x s_y}$$

- Linear Regression: $\hat{y} = mx + b$

$$m = \frac{r s_y}{s_x}, \quad b = \bar{y} - m\bar{x}$$

- Exponential Regresion: $\hat{y} = A(B^x)$

$$Y = \ln y \quad \Rightarrow \quad A = e^b, \quad B = e^m$$

- Normal Distributions

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5\left(\frac{x-\mu}{\sigma}\right)^2}$$

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

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5z^2}$$

- Central Limit Theorem

$$\mu_{\bar{x}} = \mu$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

$$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$$

- Exponential Distribution

$$\begin{aligned} f(x) &= \lambda e^{-\lambda x} \\ F(x) &= 1 - e^{-\lambda x} \\ \mu &= \lambda^{-1} \\ \sigma &= \lambda^{-1} \end{aligned}$$

- Pareto Distribution

$$\begin{aligned} f(x) &= \frac{\alpha\beta^\alpha}{x^{\alpha+1}} \\ F(x) &= 1 - \left(\frac{\beta}{x}\right)^\alpha \\ \mu &= \frac{\alpha\beta}{\alpha-1} \\ \sigma &= \sqrt{\frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)}} \end{aligned}$$

- Confidence Intervals:
  - estimating a population proportion

Confidence	α	$z_{\alpha/2}$
90%	0.10	1.645
95%	0.05	1.96
99%	0.01	2.575

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$(\hat{p} - E, \hat{p} + E)$$

  - estimating a population mean

$$df = n - 1$$

$$E = t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}$$

$$(\bar{x} - E, \bar{x} + E)$$

- estimating a population standard deviation

$$df = n - 1$$

$$\sqrt{\frac{(n-1)s^2}{\chi_R^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_L^2}}$$

- Hypothesis Tests:
  - one-sided tests

$$H_0 : \mu = \mu_0$$

$$H_a : \mu < \mu_0 \text{ or } \mu > \mu_0$$

  - two-sided tests

$$H_0 : \mu = \mu_0$$

$$H_a : \mu \neq \mu_0$$

- Goodness of Fit:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

H_o: distribution fits data well

H_a: distribution does not fit data well

- Derivatives of trig functions:

$$\begin{array}{rcl|rcl}
\frac{d}{dx}(\sin x) & = & \cos x & \frac{d}{dx}(\cos x) & = & -\sin x \\
\frac{d}{dx}(\tan x) & = & \sec^2 x & \frac{d}{dx}(\cot x) & = & -\csc^2 x \\
\frac{d}{dx}(\sec x) & = & \sec x \tan x & & & \\
\frac{d}{dx}(\csc x) & = & -\csc x \cot x & & &
\end{array}$$

- Derivatives of inverse trig functions:

$$\begin{aligned}
\frac{d}{dx}\sin^{-1} x &= \frac{1}{\sqrt{1-x^2}} \\
\frac{d}{dx}\cos^{-1} x &= \frac{-1}{\sqrt{1-x^2}} \\
\frac{d}{dx}\tan^{-1} x &= \frac{1}{1+x^2}
\end{aligned}$$

- Area between curves:

$$A = \int_a^b |f(x) - g(x)|\, dx$$

- Volume by revolution (disk/washer method):

$$V = \int_a^b \pi[r(x)]^2\, dx \quad \text{or} \quad V = \int_c^d \pi[r(y)]^2\, dy$$

- Volume by revolution (shell method):

$$V = \int_a^b 2\pi \cdot r(x) \cdot h(x)\, dx \quad \text{or} \quad V = \int_c^d 2\pi \cdot r(y) \cdot h(y)\, dy$$

- Force F

$$\begin{aligned}
F = ma &\quad \text{Newton's second law} \\
F = mg &\quad \text{force by gravity} \\
F = kx &\quad \text{Hooke's law for springs}
\end{aligned}$$

- Gravity: $g = 9.8\text{m/s}^2 = 32\text{ft/s}^2$
- Density $\rho = \dfrac{\text{mass}}{\text{volume}}$
  - water: $\rho = 1000\text{kg/m}^3 = 62.5\text{lbs./ft.}^3$
- Work $W = \int F\, dx$
- Average value of a continuous function:

$$f_{\text{avg}} = \frac{1}{b-a}\int_a^b f(x)\, dx$$

- Arc length

$$L = \int_a^b ds = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

- Area of a surface of revolution (around the x-axis)

$$S = \int_a^b 2\pi y\, ds = \int_a^b 2\pi y \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

- Center of mass (centroid) of a shape defined as the area under $f(x)$

$$\begin{aligned}
\bar{x} &= \frac{1}{A}\int_a^b x f(x)\, dx \\
\bar{y} &= \frac{1}{A}\int_a^b \frac{1}{2}[f(x)]^2\, dx
\end{aligned}$$

- Integration by parts

$$\int u\, dv = uv - \int v\, du$$

- Trigonometric Substitutions

Expression	Substitution
$\sqrt{a^2 - x^2}$	$x = a\sin\theta$
$\sqrt{a^2 + x^2}$	$x = a\tan\theta$
$\sqrt{x^2 - a^2}$	$x = a\sec\theta$

- Trapezoidal Rule $\int_a^b f(x)\,dx \approx T_n$

$$= \frac{\Delta x}{2}\left[f(x_0) + 2f(x_1) + 2f(x_2) + ... + 2f(x_{n-1}) + f(x_n)\right]$$

with $\Delta x = \dfrac{b-a}{n}$ and $x_i + ai\Delta x$ and error bound

$$|E_T| \leq \max_{x\in(a,b)} \frac{|f''(x)|(b-a)^3}{12n^2}$$

- Simpson's Rule

$$\begin{aligned}\int_a^b f(x)\,dx &\approx S_n \\ &= \frac{\Delta x}{3}[f(x_0) + 4f(x_1) + 2f(x_2) + ... \\ &\quad ... + 4f(x_{n-2}) + 2f(x_{n-1}) + f(x_n)]\end{aligned}$$

with n even, $\Delta x = \dfrac{b-a}{n}$ and $x_i + ai\Delta x$ and error bound

$$|E_S| \leq \max_{x\in(a,b)} \frac{|f^{(4)}(x)|(b-a)^5}{180n^4}$$

- p-series for improper integrals: An integral of the form

$$\int_a^\infty \frac{dx}{x^p}$$

  - *converges* to a finite value if $p > 1$
  - *diverges* toward infinity if $p \leq 1$

- Probability density function (p.d.f) over (a, b)

$$p(x) \text{ is a p.d.f. if } \int_a^b p(x)\,dx = 1$$

- Expected value, variance, and standard deviation

$$\begin{aligned}E(X) &= \int_a^b xp(x)\,dx \\ \mathrm{Var}(X) &= E(X^2) - [E(X)]^2 \\ \mathrm{SD}(X) &= \sqrt{\mathrm{Var}(X)}\end{aligned}$$

- Median: the median m of a p.d.f. is found with

$$0.5 = \int_a^m p(x)\,dx$$

- Parametric calculus (for $\alpha \leq t \leq \beta$)

$$\begin{aligned}\text{derivative:}\quad & \frac{dy}{dx} = \frac{dy/dt}{dx/dt} \\ \text{area:}\quad & A = \int_a^b y\,dx \\ \text{arc length:}\quad & L = \int_\alpha^\beta \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\,dt\end{aligned}$$

- Polar coordinates

$$x = r\cos\theta \qquad y = r\sin\theta$$
$$r^2 = x^2 + y^2 \quad \theta = \tan^{-1}\frac{y}{x}$$

- Polar calculus (for $a \leq \theta \leq b$)

$$\begin{aligned}\text{derivative:}\quad & \frac{dy}{dx} = \frac{dy/d\theta}{dx/d\theta} = \frac{\frac{dr}{d\theta}\sin\theta + r\cos\theta}{\frac{dr}{d\theta}\cos\theta - r\sin\theta} \\ \text{area:}\quad & A = \int_a^b \frac{1}{2}r^2\,d\theta \\ \text{arc length:}\quad & L = \int_a^b \sqrt{r^2 + \left(\frac{dr}{d\theta}\right)^2}\,d\theta\end{aligned}$$

- Test for (series) divergence

$$\lim_{n\to\infty} a_n \neq 0 \quad \Rightarrow \quad \sum a_n \text{ diverges}$$

- Geometric series

$$a + ar + ar^2 + ... + ar^n = \frac{a(1-r^{n+1})}{1-r}$$

and

$$\sum_{n=0}^{\infty} ar^n = a + ar + ar^2 + ...$$

converges only if $|r| < 1$ and its sum is

$$\sum_{n=1}^{\infty} ar^{n-1} = \frac{a}{1-r}$$

- Integral Test: If $f(n) = a_n$ over $(1, \infty)$, then

$$\sum a_n \text{ converges} \iff \int f(x)\,dx \text{ converges}$$

and the remainder R_n for a convergent sum $\sum a_n = s$ is bounded by

$$\int_{n+1}^{\infty} f(x)\,dx \le R_n \le \int_{n}^{\infty} f(x)\,dx$$

- Comparison Test: For sequences a_n and b_n, if $a_n \le b_n \,\forall n$, then
  1. $\sum b_n$ converges $\Rightarrow$ $\sum a_n$ converges
  2. $\sum a_n$ diverges $\Rightarrow$ $\sum b_n$ diverges
- Limit Comparison Test: Suppose that $\sum a_n$ and $\sum b_n$ are series with positive terms. If

$$\lim_{n\to\infty} \frac{a_n}{b_n} = c, \quad 0 < c < \infty$$

then either both series converge or both series diverge.

- Alternating Series Test: The alternating series $\displaystyle\sum_{n=1}^{\infty} (-1)^{n+1} b_n$ converges to a finite number s if
  1. $b_{n+1} \le b_n \quad \forall n$
  2. $\displaystyle\lim_{n\to\infty} b_n = 0$

and the remainder is bounded by

$$|R_n| = |s - s_n| \le b_{n+1}$$

- Series classification

$\sum a_n$	$\sum \lvert a_n \rvert$	classification
converges	converges	*absolute convergence*
converges	diverges	*conditional convergence*
diverges	diverges	*diverges*

- Ratio Test
  1. If $\displaystyle\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = L < 1$, then the series $\displaystyle\sum_{n=1}^{\infty} a_n$ is absolutely convergent (and therefore convergent).
  2. If $\displaystyle\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = L > 1$, or $\displaystyle\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \infty$ then the series $\displaystyle\sum_{n=1}^{\infty} a_n$ is divergent.
  3. If $\displaystyle\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = L = 1$, then further testing is needed.
- Root Test
  1. If $\displaystyle\lim_{n\to\infty} \sqrt[n]{|a_n|} = L < 1$, then the series $\displaystyle\sum_{n=1}^{\infty} a_n$ is absolutely convergent (and therefore convergent).
  2. If $\displaystyle\lim_{n\to\infty} \sqrt[n]{|a_n|} = L > 1$, or $\displaystyle\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \infty$ then the series $\displaystyle\sum_{n=1}^{\infty} a_n$ is divergent.
  3. If $\displaystyle\lim_{n\to\infty} \sqrt[n]{|a_n|} = L = 1$, then further testing is needed.
- Radius of Convergence: A power series $\displaystyle\sum_{n=0}^{\infty} a_n (x-h)^n$ centered at $x = a$ convergences in some interval $(h-R, h+R)$ where R is the *radius of convergence* (check endpoints separately).
- Taylor Series: If f has the power series representation

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n \quad |x-a| < R$$

then its coefficients are $c_n = \dfrac{f^{(n)}(a)}{n!}$

- Taylor's Inequality: If $|f^{n+1}| \le M$ for $|x-a| \le d$, then the remainder

$$|R_n(x)| \le \frac{M}{(n+1)!} |x-a|^{n+1} \quad \text{for} \quad |x-a| \le d$$

- Common Taylor series[1]

$$
\begin{aligned}
\frac{1}{1-x} &= \sum_{n=0}^{\infty} x^n \\
e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \\
\sin x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\
\cos x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \\
\tan^{-1} x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \\
(1+x)^k &= \sum_{n=0}^{\infty} \frac{k(k-1)\cdots(k-n+1)}{n!} x^n
\end{aligned}
$$

- Separable equations: A differential equation is called *separable* if the right-hand side can be written as the product of uni-variate functions

$$\frac{dy}{dx} = f(x)g(y)$$

- Euler's Method: If we wish to solve the initial-value problem

$$\frac{dy}{dt} = f(t, y), \quad y(t_0) = y_0$$

we produce Euler's Method:

$$
\begin{aligned}
t_{n+1} &= t_n + h \\
y_{n+1} &= y_n + h * f(t_n, y_n)
\end{aligned}
$$

- Logistic model: the solution to $\dfrac{dP}{dt} = kP\left(1 - \dfrac{P}{L}\right)$ is

$$P(t) = \frac{L}{1 + Ae^{-kt}} \quad \text{with} \quad A = \frac{L - P_0}{P_0}$$

[1] Note: $0! = 1$ and $\begin{pmatrix} k \\ 0 \end{pmatrix} = 1$