THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

This exam paper must not be removed from the venue

Venue ____________________

Seat Number ________

Student Number |__|__|__|__|__|__|__|__|

Family Name _____________________

First Name _____________________

School of Mathematics & Physics

EXAMINATION

Semester Two Final Examinations, 2018

STAT7203 Applied Probability and Statistics

This paper is for St Lucia Campus students.

Examination Duration: 120 minutes

Reading Time: 10 minutes

Exam Conditions:

This is a Central Examination

This is a Closed Book Examination - specified materials permitted

During reading time - write only on the rough paper provided

This examination paper will be released to the Library

Materials Permitted In The Exam Venue:

(No electronic aids are permitted e.g. laptops, phones)

Calculators - Any calculator permitted - unrestricted

One A4 sheet of handwritten or typed notes double sided is permitted

Materials To Be Supplied To Students:

None

Instructions To Students:

Additional exam materials (eg. answer booklets, rough paper) will be provided upon request.

All questions to be answered in the spaces indicated on this paper.

There are **5 questions**, worth **40 marks** in total.

Each question carries the number of marks indicated.

For Examiner Use Only

Question	Mark
1	
2	
3	
4	
5	

Total ________

1. [8 Marks] The probability generating function (pgf) of a random variable X is given by

$$G(z) = \frac{1}{3}\left(1 + z + z^2\right) .$$

(a) Determine the probability mass function (pmf) of X.

Write your answer in the box provided and show any working below.

(b) Give the mean and variance of X.

Write your answer in the box provided and show any working below.

(c) Suppose that X_1 and X_2 are independent random variables with the same distribution as X, and define $Y = X_1 + X_2$. Determine the probability mass function (pmf) of Y.

Write your answer in the box provided and show any working below.

(d) Suppose that $X_1, X_2, \ldots, X_{100}$ are independent random variables with the same distribution as X, and define $\overline{X} = \frac{1}{100}\sum_{i=1}^{100} X_i$. Use the central limit theorem (CLT) and the table of the standard Normal cumulative distribution function (cdf) to approximate $\mathbb{P}(\overline{X} \leqslant 1.1)$.

Write your answer in the box provided and show any working below.

2. [8 Marks] Consider the following model for two types of calls (standard and priority) connecting to a remote mobile-phone tower. In each one-second time slot $[0, 1), [1, 2), [2, 3), \ldots$, one standard call connects with probability 0.01 (otherwise no standard calls connect). Similarly, in each time slot, one priority call connects with probability 0.001 (otherwise no priority calls connect). Assume that calls of any type connect independently of all others.

   (a) In a 2.5 hour interval, how many calls in total are expected to connect to the mobile-phone tower?

   Write your answer in the box provided and show any working below.

(b) What is the probability that we have to wait longer than half an hour for the first priority call to connect?

Write your answer in the box provided and show any working below.

(c) Let M be the number of one-second time slots up to and including the connection of the first call of any type. Give the probability mass function (pmf) of M.

Write your answer in the box provided and show any working below.

(d) Let X be the number of one-second time-slots over the course of an entire week in which *both* a standard and priority call arrive. Use a Poisson distribution to approximate $\mathbb{P}(X = 7)$.

Write your answer in the box provided and show any working below.

3. [8 Marks] We select uniformly at random a point from the shaded (gray) region displayed in Figure 1.

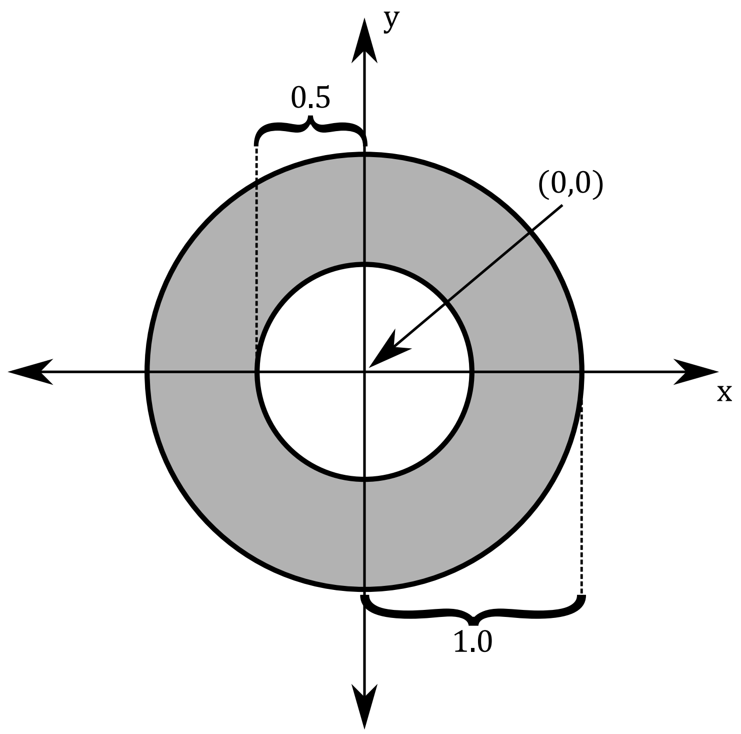

Figure 1: Annulus with unit outer radius and inner radius of 0.5.

Let X be the x-coordinate and Y the y-coordinate of the point.

(a) Determine the joint pdf of X and Y, making sure to clearly specify where it is zero.

Write your answer in the box provided.

(b) Determine the marginal pdf of Y, making sure to clearly specify where it is zero.

Write your answer in the box provided and show any working below.

(c) Determine the conditional pdf of X given $Y = 0$, making sure to clearly specify where it is zero.

Write your answer in the box provided and show any working below.

(d) Calculate $\mathbb{E}[X \mid Y = y]$ for all $y \in [-1, 1]$.

Write your answer in the box provided and show any working below.

4. [8 Marks] Capacitors Incorporated has just purchased a new machine that produces five thousand 100 micro-Farad Niobium capacitors per day, with each capacitor supposed to have exponentially distributed lifetime with mean 5 years, independently of all other capacitors produced.

   (a) Assuming that the machine is operating as it is supposed to, determine the probability that a single capacitor produced by the machine has lifetime less than or equal to 5 years.

   Write your answer in the box provided and show any working below.

(b) A quality control engineer suspects the machine is producing capacitors with shorter mean lifetimes than it is supposed to. They decide to perform a Z test for a single proportion to investigate. The engineer decides to stress-test all the capacitors produced on a certain day, and finds that the machine produced 3222 capacitors with lifetimes less than or equal to 5 years.

Write down the appropriate null hypothesis H_0, alternative hypothesis H_1, and compute the relevant test statistic.

Write your answer in the box provided and show any working below.

(c) Continuing part (b): using the table of the standard Normal cumulative distribution function, compute the p-value associated with the test.

Write your answer in the box provided and show any working below.

(d) Continuing part (c): using a significance level of $\alpha = 0.05$, is the hypothesis H_0 retained or rejected in favor of the hypothesis H_1? What can the quality control engineer conclude?

Write your answer in the box provided and show any working below.

5. [8 Marks] For a linear regression model

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i\,, \quad i = 1, \ldots, N\,,$$

where $\varepsilon_1, \ldots, \varepsilon_N \sim_{iid} \mathsf{N}(0, \sigma^2)$, denote the least-squares estimator for $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$ by $\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_0, \widehat{\beta}_1)^T$, with

$$\widehat{\beta}_1 = \frac{\left(\sum_{i=1}^N x_i Y_i\right) - N\,\bar{x}\bar{Y}}{\left(\sum_{i=1}^N x_i^2\right) - N\,\bar{x}^2} \qquad \text{and} \qquad \widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1 \bar{x}\,,$$

where as usual $\bar{x}$ denotes the average of $\{x_1, \ldots, x_N\}$.

It is known that $\widehat{\boldsymbol{\beta}}$ has a two-dimensional Normal distribution:

$$\widehat{\boldsymbol{\beta}} \sim \mathsf{N}\left(\boldsymbol{\beta}, \frac{\sigma^2}{\left(\sum_{i=1}^N x_i^2\right) - N\,\bar{x}^2} \begin{pmatrix} \frac{1}{N}\sum_{i=1}^N x_i^2 & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix}\right).$$

Suppose we construct a linear regression model for the time it takes (Y, in seconds) to write a file to a new solid-state hard disk with file size (x, in gigabytes) as the explanatory variable, with $\sigma = 0.5$ seconds.

Out of $N = 20$ file transfers, the following summary statistics are collected.

Summary Statistic	Value	Units
$\sum_{i=1}^N x_i$	58.83	gigabytes
$\sum_{i=1}^N x_i^2$	200.99	gigabytes^2
$\sum_{i=1}^N y_i$	126.97	seconds
$\sum_{i=1}^N x_i y_i$	424.26	gigabyte-seconds

(a) Using the summary statistics, compute the value for the estimates $\widehat{\beta}_0$ and $\widehat{\beta}_1$.

Write your answer in the box provided and show any working on the next page.

Working space only.

(b) Using the value for $\widehat{\beta}_0$ computed in (a) and the given Normal distribution for $\widehat{\beta}_0$, determine a 95% numerical confidence interval for $\widehat{\beta}_0$.

Write your answer in the box provided and show any working below.

(c) Using the values for $\widehat{\beta}_0$ and $\widehat{\beta}_1$ computed in (a), what is the estimated time in seconds it takes to write a file of 2.2 gigabytes?

Write your answer in the box provided and show any working below.

(d) The predicted time (in seconds) it takes to write a new file of x gigabytes is given by

$$\widehat{Y(x)} = \widehat{\beta}_0 + \widehat{\beta}_1 \, x + \varepsilon \, ,$$

where $\varepsilon \sim \mathsf{N}(0, \sigma^2)$. It is known that $\widehat{Y(x)}$ will be normally distributed. Using the given two-dimensional Normal distribution for $\widehat{\boldsymbol{\beta}}$, determine the mean and variance of $\widehat{Y(x)}$ (both as a function of x).

Write your answer in the box provided and show any working below.

Working space only.

(e) Using the values for $\widehat{\beta}_0$ and $\widehat{\beta}_1$ computed in (a) and the normal distribution found in (c), determine a 95% numerical prediction interval for the time (in seconds) it takes to write a new file of 2.2 gigabytes.

Write your answer in the box provided and show any working below.

Total [40 Marks]

Standard Normal distribution

This table gives the cumulative distribution function (cdf) Φ of a $\mathsf{N}(0,1)$-distributed random variable Z.

$$\Phi(z) = \mathbb{P}(Z \leqslant z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2}\, dx.$$

The last column gives the probability density function (pdf) φ of the $\mathsf{N}(0,1)$-distribution

$$\varphi(z) = \frac{1}{\sqrt{2\pi}}\, e^{-z^2/2}.$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09	$\varphi(z)$
0.0	5000	5040	5080	5120	5160	5199	5239	5279	5319	5359	0.3989
0.1	5398	5438	5478	5517	5557	5596	5636	5675	5714	5753	0.3970
0.2	5793	5832	5871	5910	5948	5987	6026	6064	6103	6141	0.3910
0.3	6179	6217	6255	6293	6331	6368	6406	6443	6480	6517	0.3814
0.4	6554	6591	6628	6664	6700	6736	6772	6808	6844	6879	0.3683
0.5	6915	6950	6985	7019	7054	7088	7123	7157	7190	7224	0.3521
0.6	7257	7291	7324	7357	7389	7422	7454	7486	7517	7549	0.3332
0.7	7580	7611	7642	7673	7704	7734	7764	7794	7823	7852	0.3123
0.8	7881	7910	7939	7967	7995	8023	8051	8078	8106	8133	0.2897
0.9	8159	8186	8212	8238	8264	8289	8315	8340	8365	8389	0.2661
1.0	8413	8438	8461	8485	8508	8531	8554	8577	8599	8621	0.2420
1.1	8643	8665	8686	8708	8729	8749	8770	8790	8810	8830	0.2179
1.2	8849	8869	8888	8907	8925	8944	8962	8980	8997	9015	0.1942
1.3	9032	9049	9066	9082	9099	9115	9131	9147	9162	9177	0.1714
1.4	9192	9207	9222	9236	9251	9265	9279	9292	9306	9319	0.1497
1.5	9332	9345	9357	9370	9382	9394	9406	9418	9429	9441	0.1295
1.6	9452	9463	9474	9484	9495	9505	9515	9525	9535	9545	0.1109
1.7	9554	9564	9573	9582	9591	9599	9608	9616	9625	9633	0.0940
1.8	9641	9649	9656	9664	9671	9678	9686	9693	9699	9706	0.0790
1.9	9713	9719	9726	9732	9738	9744	9750	9756	9761	9767	0.0656
2.0	9772	9778	9783	9788	9793	9798	9803	9808	9812	9817	0.0540
2.1	9821	9826	9830	9834	9838	9842	9846	9850	9854	9857	0.0440
2.2	9861	9864	9868	9871	9875	9878	9881	9884	9887	9890	0.0355
2.3	9893	9896	9898	9901	9904	9906	9909	9911	9913	9916	0.0283
2.4	9918	9920	9922	9925	9927	9929	9931	9932	9934	9936	0.0224
2.5	9938	9940	9941	9943	9945	9946	9948	9949	9951	9952	0.0175
2.6	9953	9955	9956	9957	9959	9960	9961	9962	9963	9964	0.0136
2.7	9965	9966	9967	9968	9969	9970	9971	9972	9973	9974	0.0104
2.8	9974	9975	9976	9977	9977	9978	9979	9979	9980	9981	0.0079
2.9	9981	9982	9982	9983	9984	9984	9985	9985	9986	9986	0.0060
3.0	9987	9987	9987	9988	9988	9989	9989	9989	9990	9990	0.0044
3.1	9990	9991	9991	9991	9992	9992	9992	9992	9993	9993	0.0033
3.2	9993	9993	9994	9994	9994	9994	9994	9995	9995	9995	0.0024
3.3	9995	9995	9995	9996	9996	9996	9996	9996	9996	9997	0.0017
3.4	9997	9997	9997	9997	9997	9997	9997	9997	9997	9998	0.0012
3.5	9998	9998	9998	9998	9998	9998	9998	9998	9998	9998	0.0009
3.6	9998	9998	9999	9999	9999	9999	9999	9999	9999	9999	0.0006

Example: $\Phi(1.65) = \mathbb{P}(Z \leqslant 1.65) = 0.9505$

Summary of Formulas

1. **Sum rule:** $\boxed{\mathbb{P}(\bigcup_i A_i) = \sum_i \mathbb{P}(A_i)}$, when $A_1, A_2, \dots$ are disjoint.

2. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

3. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

4. **Cdf** of X: $F(x) = \mathbb{P}(X \leqslant x),\ x \in \mathbb{R}$.

5. **Pmf** of X: (discrete r.v.) $f(x) = \mathbb{P}(X = x)$.

6. **Pdf** of X: (continuous r.v.) $f(x) = F'(x)$.

7. For a discrete r.v. X: $\boxed{\mathbb{P}(X \in B) = \sum_{x \in B} \mathbb{P}(X = x).}$

8. For a continuous r.v. X with pdf f: $\boxed{\mathbb{P}(X \in B) = \int_B f(x)\, dx.}$

9. In particular (continuous), $F(x) = \int_{-\infty}^{x} f(u)\, du$.

10. Similar results 7-8 hold for random vectors, e.g. $\mathbb{P}((X,Y) \in B) = \iint_B f_{X,Y}(x,y)\, dx\, dy$.

11. Marginal from joint pdf: $f_X(x) = \int f_{X,Y}(x,y)\, dy$.

12. **Important discrete distributions**:

Distr.	pmf	$x \in$
$\mathsf{Ber}(p)$	$p^x(1-p)^{1-x}$	$\{0,1\}$
$\mathsf{Bin}(n,p)$	$\binom{n}{x} p^x (1-p)^{n-x}$	$\{0,1,\dots,n\}$
$\mathsf{Poi}(\lambda)$	$\mathrm{e}^{-\lambda} \frac{\lambda^x}{x!}$	$\{0,1,\dots\}$
$\mathsf{Geom}(p)$	$p(1-p)^{x-1}$	$\{1,2,\dots\}$

13. **Important continuous distributions**:

Distr.	pdf	$x \in$
$\mathsf{U}[a,b]$	$\frac{1}{b-a}$	$[a,b]$
$\mathsf{Exp}(\lambda)$	$\lambda\, \mathrm{e}^{-\lambda x}$	$\mathbb{R}_+$
$\mathsf{Gamma}(\alpha,\lambda)$	$\frac{\lambda^\alpha x^{\alpha-1} \mathrm{e}^{-\lambda x}}{\Gamma(\alpha)}$	$\mathbb{R}_+$
$\mathsf{N}(\mu,\sigma^2)$	$\frac{1}{\sigma\sqrt{2\pi}}\, \mathrm{e}^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	$\mathbb{R}$

14. **Conditional probability:** $\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$.

15. **Law of total probability**: $\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \mid B_i)\, \mathbb{P}(B_i)$, with $B_1, B_2, \dots, B_n$ a partition of Ω.

16. **Bayes' Rule:** $\mathbb{P}(B_j|A) = \frac{\mathbb{P}(B_j)\,\mathbb{P}(A|B_j)}{\sum_{i=1}^{n} \mathbb{P}(B_i)\,\mathbb{P}(A|B_i)}$.

17. **Product rule**: $\boxed{\mathbb{P}(A_1 \cdots A_n) = \mathbb{P}(A_1)\,\mathbb{P}(A_2 \mid A_1) \cdots \mathbb{P}(A_n \mid A_1 \cdots A_{n-1}).}$

18. **Memoryless property** (Exp and Geom distribution): $\mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t),\ \forall s, t$.

19. **Independent events**: $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.

20. **Independent r.v.'s:** (discrete) $\boxed{\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \prod_{k=1}^{n} \mathbb{P}(X_k = x_k)\ .}$

21. **Independent r.v.'s:** (continuous) $\boxed{f_{X_1,\dots,X_n}(x_1,\dots,x_n) = \prod_{k=1}^{n} f_{X_k}(x_k)\ .}$

22. **Expectation** (discr.): $\mathbb{E}X = \sum_x x\, \mathbb{P}(X = x)$.

23. (of function) $\mathbb{E}\, g(X) = \sum_x g(x)\, \mathbb{P}(X = x)$.

24. **Expectation** (cont.): $\mathbb{E}X = \int x\, f(x)\, dx$.

25. (of function) $\mathbb{E}\, g(X) = \int g(x) f(x)\, dx$,

26. Similar results 22–25 hold for random vectors.

27. **Expected sum :** $\mathbb{E}(aX + bY) = a\, \mathbb{E}X + b\, \mathbb{E}Y$.

28. **Expected product** (only if X, Y independent): $\mathbb{E}[X\, Y] = \mathbb{E}X\, \mathbb{E}Y$.

29. **Markov inequality:** $\mathbb{P}(X \geqslant x) \leqslant \frac{\mathbb{E}X}{x}$.

30. $\mathbb{E}X$ **and Var**(X) **for various distributions**:

	$\mathbb{E}X$	$\mathrm{Var}(X)$
$\mathsf{Ber}(p)$	p	$p(1-p)$
$\mathsf{Bin}(n,p)$	np	$np(1-p)$
$\mathsf{Geom}(p)$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$\mathsf{Poi}(\lambda)$	λ	λ
$\mathsf{U}(a,b)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$\mathsf{Exp}(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$\mathsf{Gamma}(\alpha,\lambda)$	$\frac{\alpha}{\lambda}$	$\frac{\alpha}{\lambda^2}$
$\mathsf{N}(\mu,\sigma^2)$	μ	σ^2

31. n**-th moment:** $\mathbb{E}X^n$.

32. **Covariance:** $\mathrm{cov}(X,Y) = \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y)$.

33. **Properties of Var and Cov**:

$\mathrm{Var}(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2$.
$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$.
$\mathrm{cov}(X,Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y$.
$\mathrm{cov}(X,Y) = \mathrm{cov}(Y,X)$.
$\mathrm{cov}(aX + bY, Z) = a\,\mathrm{cov}(X,Z) + b\,\mathrm{cov}(Y,Z)$.
$\mathrm{cov}(X,X) = \mathrm{Var}(X)$.
$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{cov}(X,Y)$.
X and Y independent $\Longrightarrow \mathrm{cov}(X,Y) = 0$.

34. **Probability Generating Function (PGF)**: $G(z) := \mathbb{E}\, z^N = \sum_{n=0}^{\infty} \mathbb{P}(N = n) z^n, \qquad |z| < 1$.

35. **PGFs for various distributions:**

$\mathsf{Ber}(p)$	$1 - p + zp$
$\mathsf{Bin}(n,p)$	$(1 - p + zp)^n$
$\mathsf{Geom}(p)$	$\frac{z\,p}{1 - z\,(1-p)}$
$\mathsf{Poi}(\lambda)$	$\mathrm{e}^{-\lambda(1-z)}$

36. $\mathbb{P}(N = n) = \frac{1}{n!}\, G^{(n)}(0)$. (n-th derivative, at 0)

37. $\mathbb{E}N = G'(1)$

38. $\mathrm{Var}(N) = G''(1) + G'(1) - (G'(1))^2$.

39. **Moment Generating Function (MGF)**: $M(s) = \mathbb{E}\,e^{sX} = \int_{-\infty}^{\infty} e^{sx} f(x)\,dx$, $s \in I \subset \mathbb{R}$, for r.v.'s X for which all moments exist.

40. **MGFs for various distributions:**

$\mathsf{U}(a,b)$	$\frac{e^{bs}-e^{as}}{s(b-a)}$
$\mathsf{Gamma}(\alpha,\lambda)$	$\left(\frac{\lambda}{\lambda-s}\right)^{\alpha}$
$\mathsf{N}(\mu,\sigma^2)$	$e^{s\mu+\sigma^2 s^2/2}$

41. **Moment property**: $\mathbb{E}X^n = M^{(n)}(0)$.

42. $M_{X+Y}(t) = M_X(t)\,M_Y(t)$, $\forall t$, if X, Y independent.

43. If $X_i \sim \mathsf{N}(\mu_i, \sigma_i^2)$, $i = 1, 2, \ldots, n$ (independent), then $a + \sum_{i=1}^n b_i X_i \sim \mathsf{N}\left(a + \sum_{i=1}^n b_i \mu_i,\ \sum_{i=1}^n b_i^2 \sigma_i^2\right)$.

44. **Conditional pmf/pdf**

$$f_{Y\,|\,X}(y\,|\,x) := \frac{f_{X,Y}(x,y)}{f_X(x)}, \quad y \in \mathbb{R}.$$

45. The corresponding **conditional expectation** (discrete case): $\mathbb{E}[Y\,|\,X=x] = \sum_y y\, f_{Y\,|\,X}(y\,|\,x)$.

46. **Linear transformation**: $f_{\boldsymbol{Z}}(\boldsymbol{z}) = \frac{f_{\boldsymbol{X}}(A^{-1}\boldsymbol{z})}{|A|}$.

47. **General transformation**: $f_{\boldsymbol{Z}}(\boldsymbol{z}) = \dfrac{f_{\boldsymbol{X}}(\mathbf{x})}{|J_{\mathbf{x}}(g)|}$, with $\mathbf{x} = g^{-1}(\boldsymbol{z})$, where $|J_{\mathbf{x}}(g)|$ is the Jacobian of g evaluated at $\boldsymbol{x}$.

48. Pdf of the **multivariate Normal** distribution:

$$f_{\boldsymbol{Z}}(\boldsymbol{z}) = \frac{1}{\sqrt{(2\pi)^n\,|\Sigma|}}\, e^{-\frac{1}{2}(\boldsymbol{z}-\boldsymbol{\mu})^T \Sigma^{-1}(\boldsymbol{z}-\boldsymbol{\mu})} .$$

Σ is the covariance matrix, and $\boldsymbol{\mu}$ the mean vector.

49. If $\mathbf{X}$ is a column vector with independent $\mathsf{N}(0,1)$ components, and B is a matrix with $\Sigma = BB^T$ (such a B can always be found), then $\mathbf{Z} = \boldsymbol{\mu} + B\mathbf{X}$ has a multivariate Normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ.

50. **Weak Law of Large Numbers:**

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) = 0, \quad \forall \varepsilon .$$

51. **Strong Law of Large Numbers**:

$$\mathbb{P}\left(\lim_{n\to\infty} \frac{S_n}{n} = \mu\right) = 1,$$

as $n \to \infty$.

52. **Central Limit Theorem**:

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leqslant x\right) = \Phi(x),$$

where Φ is the cdf of the standard Normal distribution.

53. **Normal Approximation to Binomial:** If $X \sim \mathsf{Bin}(n,p)$, then, for large n, $\mathbb{P}(X \leqslant k) \approx \mathbb{P}(Y \leqslant k)$, where $Y \sim \mathsf{N}(np, np(1-p))$.

Other Mathematical Formulas

1. Factorial. $n! = n\,(n-1)\,(n-2)\cdots 1$. Gives the number of *permutations* (orderings) of $\{1,\ldots,n\}$.

2. Binomial coefficient. $\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$. Gives the number *combinations* (no order) of k different numbers from $\{1,\ldots,n\}$.

3. Newton's binomial theorem: $(a+b)^n = \sum_{k=0}^n a^k\, b^{n-k}$.

4. Geometric sum: $1 + a + a^2 + \cdots + a^n = \frac{1-a^{n+1}}{1-a}$ $(a \neq 1)$.
If $|a| < 1$ then $1 + a + a^2 + \cdots = \frac{1}{1-a}$.

5. Logarithms:
   (a) $\log(x\,y) = \log x + \log y$.
   (b) $e^{\log x} = x$.

6. Exponential:
   (a) $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$.
   (b) $e^x = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n$.
   (c) $e^{x+y} = e^x\, e^y$.

7. Differentiation:
   (a) $(f+g)' = f' + g'$
   (b) $(fg)' = f'g + fg'$
   (c) $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$
   (d) $\frac{d}{dx}x^n = n\,x^{n-1}$
   (e) $\frac{d}{dx}e^x = e^x$
   (f) $\frac{d}{dx}\log(x) = \frac{1}{x}$

8. Chain rule: $(f(g(x)))' = f'(g(x))\,g'(x)$.

9. Integration: $\int_a^b f(x)\,dx = [F(x)]_a^b = F(b) - F(a)$, where $F' = f$.

10. Integration by parts: $\int_a^b f(x)\,G(x)\,dx = [F(x)\,G(x)]_a^b - \int_a^b F(x)\,g(x)\,dx$. (Here $F' = f$ and $G' = f$.)

11. Jacobian: Let $\mathbf{x} = (x_1,\ldots,x_n)$ be an n-dimensional vector, and $g(\mathbf{x}) = (g_1(\mathbf{x}),\ldots,g_n(\mathbf{x}))$ be a function from $\mathbb{R}^n$ to $\mathbb{R}^n$. The *matrix of Jacobi* is the matrix of partial derivatives: $(\partial g_i/\partial x_j)$. The corresponding determinant is called the *Jacobian*. In the neighbourhood of any fixed point, g behaves like a *linear transformation* specified by the matrix of Jacobi at that point.

12. Γ function: $\Gamma(\alpha) = \int_0^\infty u^{\alpha-1}\,e^{-u}\,du$, $\alpha > 0$. $\Gamma(\alpha+1) = \alpha\,\Gamma(\alpha)$, for $\alpha \in \mathbb{R}_+$. $\Gamma(n) = (n-1)!$ for $n = 1, 2, \ldots$. $\Gamma(1/2) = \sqrt{\pi}$.