Mr. G's little booklet on

Probability and Statistics

Issue 5.0 10/16

rg

Mr. G's Little Booklets are

1 **Symbols and Definitions**

2 **Circular Functions**

3 **Hyperbolic Functions**

4 **Complex Numbers**

5 **Calculus**

6 **Series**

7 **Venn Diagrams**

8 **Logic and Propositional Calculus**

9 **Vectors and Matrices**

10 **Probability**

11 **Laplace and Fourier Transforms**

12 **Miscellaneous Aspects of Mathematics**

13 **Statistical Tables**

14 **Trigonometric and Logarithmic Tables**

15 **Investigations - General**

16 **Investigations - Number**

rg

Probability

$P(A) \geq 0$		*for every event A*	*Axiom 1*
$P(S) = 1$		*for the CERTAIN event S*	*Axiom 2*
$P(A) = {}^{n(A)}/_{n(U)}$			*Axiom 3*
$P(A \cup B) = P(A) + P(B)$		*for mutually exclusive events A and B.*	*corollary 1*
$P(\varnothing) = 0$			*Theorem 1*
$P(A \cap B') = P(A) - P(A \cap B)$		*because* $A \cap B$ & $A \cap B'$ *are disjoint*	*Theorem 2*
$P(B') = 1 - P(B)$		*by putting A = S in theorem 2*	*Theorem 3*
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$			*Theorem 4*

Independent Events

if $P(A \cap B) = P(A) \times P(B)$ *then events A and B are independent*

If events A *and* B *are independent then so are events* A *and* B' — *Theorem 5*

If events A *and* B *are independent with non zero probabilities then* $A \cap B \neq \varnothing$ — *Theorem 6*

Conditional Probability and Bayes Theorem

define $P(A|B)$ = *conditional probability of A given that B has already occurred*

Let $P(A \cap B) \quad P(A|B) \times P(B)$

Hence $P(A|B) = {}^{P(A \cap B)}/_{P(B)}$

Also $P(B|A) = {}^{P(B \cap A)}/_{P(A)}$ *exchanging A and B*

${}^{P(A|B)}/_{P(B|A)} = {}^{P(A)}/_{P(B)}$ *by commutative law* $A \cap B = B \cap A$

So $P(A|B) = {}^{P(A)\,P(B|A)}/_{P(B)}$

So $P(A|B) = {}^{P(A)\,P(B|A)}/_{\{P(A \cap B) + P(A' \cap B)\}}$

So $P(A|B) = {}^{P(A)\,P(B|A)}/_{P(A)\,P(B|A) + P(A')\,P(B|A')}$ *Bayes*

Note the derivation of the useful relationship $P(B) = P(B|A) \times P(A) + P(B|A') \times P(A')$

This is the probability that event A which has occurred is a result of cause B.

If B *is the new evidence then* ${}^{P(B|A)}/_{P(B)}$ *is called the Likelihood Ratio*

nb if A *and* B *are* *<u>independent</u>* *then* $P(A|B) = P(A)$

nb if A *and* B *are* *<u>mutually exclusive</u>* *then* $P(A|B) = 0$

Notes

mutually exclusive events may be viewed as A **OR** B **OR** C *etc.*

inedependent events may be viewed as A **AND** (then) B **AND** (then) C etc.

Strict Permutations *(cf general permutations)* $^nP_r = {}^{n!}/_{(n-r)!}$

Consider Set S with elements {C,D,E,F} n(A) = 4

A permutation is a sequence containing each element of a set once and only once.

For set size n the number of permutations is n! (in this case 4 x 3 x 2 x 1 = 24)

C	D	E	F	D	C	E	F	E	C	D	F	F	C	D	E
C	D	F	E	D	C	F	E	E	C	F	D	F	C	E	D
C	E	D	F	D	E	C	F	E	D	C	F	F	D	C	E
C	E	F	D	D	E	F	C	E	D	F	C	F	D	E	C
C	F	D	E	D	F	C	E	E	F	C	D	F	E	C	D
C	F	E	D	D	F	E	C	E	F	D	C	F	E	D	C

Combinations $^nC_r = {}^{n!}/_{r!(n-r)!}$ *write* $\binom{n}{r}$

A combination is an unordered subset of a specified number of distinct elements

$^4C_1$	=4	$^4C_2$	=6		$^4C_3$	=4	$^4C_4$	=1
{C}	{E}	{C,D}	{C,F}	{D,F}	{C,D,E}	{D,E,F}	{C,D,E,F}	
{D}	{F}	{C,E}	{D,E}	{E,F}	{C,D,F}	{C,E,F}		

Power Sets

For the set S the powerset of S is P(S) and is the set of all subsets including the null set.

It is thus the union of all the combinations of all sizes plus the null set { }.

The power set contains 2^n elements. For n(S) =4 there are 2^4 =16 subsets.

{C}	{E}	{C,D}	{C,F}	{D,F}	{C,D,E}	{D,E,F}	{C,D,E,F}
{D}	{F}	{C,E}	{D,E}	{E,F}	{C,D,F}	{C,E,F}	plus { }

Cardinality (Size)

It can proved that the powerset of a set has a higher cardinality than the original set.

The proof is valid for finite sets <u>*and also infinite*</u> *sets.*

If we start with $\aleph_0 = \infty$ *then* $2^\infty = \aleph_1$ *where* $\aleph_1 > \infty$

The number of points on a line, n{ X }*, is denoted* C *where* C > $\aleph_0$

It would be neat to suppose that $\aleph_1$ = C *but this is indeterminate within set theory.*

That is you can assume it to be either true or false and neither will lead to a contradiction.

Binomial Distribution

$$\{ p + (1-p) \}^n = \sum \binom{n}{r} p^r (1-p)^{n-r}$$

$$\textit{where } \binom{n}{r} = {}^nC_r$$

$$\binom{n}{r} = \binom{n}{n-r} \qquad \textit{Theorem 1}$$

$$\binom{n}{r-1} + \binom{n}{r} = \binom{n+1}{r} \qquad \textit{Theorem 2}$$

$$\binom{n}{0}+\binom{n}{1}+\binom{n}{2}\ldots\binom{n}{n} = 2^n \qquad \textit{Theorem 3}$$

$$P(\text{r successes from n trials}) = \binom{n}{r} p^r q^{n-r} \qquad \textit{where } p + q = 1$$

If $X \sim B(n,p)$ then

$$E(X) = np$$

$$Var(X) = npq$$

Poisson Distribution

$$P(X = r) = e^{-\lambda} \lambda^r / r!$$

where X is no. events in given interval and λ is mean no. events in same interval

$$E(X) = \lambda$$

$$Var(X) = \lambda$$

Poisson approximates to the Binomial for n large and p small

Normal Distribution

$$\Phi(z) = \{^1/_{\sqrt{(2\pi)}}\} \int_{-\infty}^{z} e^{-\frac{1}{2}t^2} dt$$

$$= \{1 + \text{erf}\,(^z/_{\sqrt{2}})\}$$

$$E(X) = \mu$$

$$Var(X) = \sigma^2$$

given $X \sim N(\mu,\sigma^2)$

$$Z = {}^{(X-\mu)}/_{\sigma} \sim N(0,1)$$

Notes

sech² *is a bell shaped distribution occurring in the natural world.*

Properties of Binomial Coefficients

$\binom{n}{0}$	+	$\binom{n}{1}$	+	$\binom{n}{2}$	+	…	+	$\binom{n}{n}$	=	2^n
1	+	4	+	6	+	4	+	1	=	$2^4 = 16$
$\binom{n}{0}$	−	$\binom{n}{1}$	+	$\binom{n}{2}$	−	…	$(^-1)^n$	$\binom{n}{n}$	=	0
1	−	4	+	6	−	4	+	1	=	0
$\binom{n}{n}$	+	$\binom{n+1}{n}$	+	$\binom{n+2}{n}$	+	…	+	$\binom{n+m}{n}$	=	$\binom{n+m+1}{n+1}$
1	+	5	+	15	+	35	+	70	=	$\binom{9}{5} = 126$
$\binom{n}{0}$	+	$\binom{n}{2}$	+	$\binom{n}{4}$	+	$\binom{n}{8}$	+	…	=	2^{n-1}
1	+	28	+	70	+	28	+	1	=	$2^7 = 128$
$\binom{n}{1}$	+	$\binom{n}{3}$	+	$\binom{n}{5}$	+	$\binom{n}{9}$	+	…	=	2^{n-1}
9	+	84	+	126	+	36	+	1	=	$2^8 = 256$
$\binom{n}{0}^2$	+	$\binom{n}{1}^2$	+	$\binom{n}{2}^2$	+	…	+	$\binom{n}{n}^2$	=	$\binom{2n}{n}$
1	+	16	+	36	+	16	+	1	=	$\binom{8}{4} = 70$
$\binom{m}{0}\binom{n}{p}$	+	$\binom{m}{1}\binom{n}{p-1}$	+	$\binom{m}{2}\binom{n}{p-2}$	+	$\binom{m}{4}\binom{n}{p-4}$	+	$\binom{m}{p}\binom{n}{0}$	=	$\binom{m+n}{p}$
1×15	+	7×20	+	21×15	+	35×6	+	35×1	=	$\binom{7+6}{4}=715$
$(1)\binom{n}{1}$	+	$(2)\binom{n}{2}$	+	$(3)\binom{n}{3}$	+	…	+	$(n)\binom{n}{n}$	=	$n2^{n-1}$
1×5	+	2×10	+	3×10	+	4×5	+	5×1	=	$5\times2^4=80$
$(1)\binom{n}{1}$	−	$(2)\binom{n}{2}$	+		+	…	$(^-1)^{n+1}$	$(n)\binom{n}{n}$	=	0
1×5	−	2×10	+	3×10	−	4×5	+	5×1	=	0

Notes

To appreciate fully the interpretation of these properties trace them out on Pascal's Triangle

Pascal's Triangle

Σ	2^1	2^2	2^3	2^4	2^5	2^6	2^7	2^8	2^9	2^{10}	1
										1	
									1		11
								1		10	
							1		9		55
						1		8		45	
					1		7		36		165
				1		6		28		120	
			1		5		21		84		330
		1		4		15		56		210	
	1		3		10		35		126		462
1		2		6		20		70		252	
	1		3		10		35		126		462
		1		4		15		56		210	
			1		5		21		84		330
				1		6		28		120	
					1		7		36		165
						1		8		45	
							1		9		55
								1		10	
									1		11
										1	
											1

Integers

Triangulars

Tetrahedrals

Pentatopes

Notes

Pascal's Triangle as developed by Isaac Newton

n =	-5	1	-5	15	-35	70	-126	210	-330	495	-715
n =	-4	1	-4	10	-20	35	-56	84	-120	165	-220
n =	-3	1	-3	6	-10	15	-21	28	-36	45	-55
n =	-2	1	-2	3	-4	5	-6	7	-8	9	-10
n =	-1	1	-1	1	-1	1	-1	1	-1	1	-1
n =	0	1	0	0	0	0	0	0	0	0	0
n =	½	1	$^{1}/_{2}$	$^{-1}/_{8}$	$^{1}/_{16}$	$^{-5}/_{128}$	$^{7}/_{256}$	$^{-21}/_{1024}$	$^{33}/_{2048}$	$^{-429}/_{32768}$	$^{715}/_{65536}$
n =	1	1	1	0	0	0	0	0	0	0	0
n =	2	1	2	1	0	0	0	0	0	0	0
n =	3	1	3	3	1	0	0	0	0	0	0
n =	4	1	4	6	4	1	0	0	0	0	0
n =	5	1	5	10	10	5	1	0	0	0	0
n =	6	1	6	15	20	15	6	1	0	0	0
n =	7	1	7	21	35	35	21	7	1	0	0
n =	8	1	8	28	56	70	56	28	8	1	0
n =	9	1	9	36	84	126	126	84	36	9	1
n =	10	1	10	45	120	210	252	210	120	45	10
n =	11	1	11	55	165	330	462	462	330	165	55
n =	12	1	12	66	220	495	792	924	792	495	220
n =	13	1	13	78	286	715	1287	1716	1716	1287	715
n =	14	1	14	91	364	1001	2002	3003	3432	3003	2002
n =	15	1	15	105	455	1365	3003	5005	6435	6435	5005

Notes

Newton extended Pascal's triangle into negative values.

He also interpolated between integers to derive terms for fractional values of n.

We can derive terms from the binomial expansion $1 + nx + {}^{n(n-1)x^2}/_{2!} + {}^{n(n-1)(n-2)x^3}/_{3!} \ldots$

Definition of Moments of a Distribution

Sample variate X *has a distribution* P(x), *known population mean* μ *and variance* σ^2

	Population	**Sample**	**Unbiased estimate**
mean	μ	$\tilde{x} = {}^{\Sigma fx}/_{\Sigma f}$	$\tilde{x}$
variance	$\sigma^2 = \langle (X - \mu)^2 \rangle$	$s_n{}^2$	$s_{n-1}{}^2$
discrete	$\sum (x_i)\ (x_i - \mu)^2$	${}^1/_n \sum (x_i - \tilde{x})^2$ *(mean square deviation msd)*	${}^1/_{n-1} \sum (x_i - \tilde{x})^2$ ‡
continuous	$\int P(x)(x - \mu)^2\, dx$		
standard deviation	$\sigma = {}^+\sqrt{\sigma^2}$	$s_n = {}^+\sqrt{s_n{}^2}$	$s_{n-1} = {}^+\sqrt{s_{n-1}{}^2}$

Correlation

$r = S_{xy} / (S_{xx} S_{yy})$ *pearson product moment corr. coeff.*

where $S_{xy} = \sum (x_i - \tilde{x})(y_i - \tilde{y})$

where $S_{xx} = \sum (x_i - \tilde{x})^2$

$= \sum (x_i{}^2) - {}^{(\Sigma xi)^2}/_n$

where $S_{yy} = \sum (y_i - \tilde{y})^2$

$= \sum (y_i{}^2) - {}^{(\Sigma yi)^2}/_n$

Regression

$y = ax + b$ *regression line y on x*

$a = S_{xy} / S_{xx}$

$b = \tilde{y} - a\tilde{x}$

Notes

† *This is the closest I can get to writing x bar*

‡ *This is the* κ_2 *statistic given by* $\kappa_2 = {}^n/_{n-1}\ s_n{}^2$

Determination (r^2) and Correlation (r)

		r^2	=	0.00	*no correlation*	
		r	=	0.00	*no correlation*	
0.00	<	r^2	≤	0.25	*very weak correlation*	
0.00	<	r	≤	0.50	*very weak correlation*	
0.25	<	r^2	≤	0.50	*weak correlation*	
0.50	<	r	≤	0.71	*weak correlation*	†
0.50	<	r^2	≤	0.75	*moderate correlation*	
0.71	<	r	≤	0.87	*moderate correlation*	†
0.75	<	r^2	≤	0.90	*strong correlation*	
0.87	<	r	≤	0.95	*strong correlation*	
0.90	<	r^2	≤	1.00	*very strong correlation*	
0.95	<	r	≤	1.00	*very strong correlation*	
		r^2	=	1.00	*perfect correlation*	
		r	=	1.00	*perfect correlation*	

Notes

† *These values shouldn't be taken too literally.*

Note that correlation does not imply causation

The coefficient of determination r^2 is a measure of the strength of association between two variables, X and Y.

For linear regression, it is the proportion explained by the conjectured model.

The correlation coefficient r is a measure of the degree of linearity between two variables

In the case of perfect linear correlation we have $r^2 = r^2$

r^2, *the coefficient of determination, is a thing all in its own right*

r^2 *is r, the correlation coefficient, squared.*

Students often get confused between the two expressions - no idea why.

Greek Alphabet				Principle/Simplest Use	English	Type
alpha	Α	*not used*	α	*first root of quadratic*	a	**a**
beta	Β	*Beta function*	β	*second root of quadratic*	b	**b**
gamma	Γ	*Gamma function*	γ	*Euler's constant*	g	**g**
delta	Δ	*Difference operator*	δ	*small increment*	d	**d**
epsilon	Ε	*not used*	ε	*error*	short e	**e**
zeta	Ζ	*not used*	ζ	*Riemann zeta function*	z	**z**
eta	Η	*not used*	η	*efficiency*	long e	**h**
theta	Θ	*asymp. tight bound*	θ	*angle*	th	**q**
iota	Ι	*not used*	ι	*imaginary unit*	i	**i**
kappa	Κ	*not used*	κ	*curvature*	k	**k**
lambda	Λ	*diag. matrix eigen-values*	λ	*failure rate*	l	**l**
mu	Μ	*not used*	μ	*population mean*	m	**m**
nu	Ν	*not used*	ν	*poisson ratio*	n	**n**
xi	Ξ	*grand canonical ensemble*	ξ	*damping coefficient*	x	**x**
omicron	Ο	*limiting behaviour function*	ο	*generally not used*	short o	**o**
pi	Π	*Product operator*	π	*ratio* $^{c}/_{d}$ *circle*	p	**p**
rho	Ρ	*not used*	ρ	*correlation coefficient*	r	**r**
sigma	Σ	*summation*	σ	*standard deviation*	s	**s**
tau	Τ	*not used*	τ	*mean lifetime*	t	**t**
upsilon	Υ	*Bessel function*	υ	*generally not used*	u	**u**
phi	Φ	*cumulative function*	ϕ	*golden ratio*	ph	**f**
phi (alt.)	ϑ	*not used*	φ	**normal function** **scalar potential**	ph	**j**
chi	Χ	*probability function*	χ^2	*chi-squared prob.function*	ch	**c**
psi	Ψ	*not used*	ψ	*wave function*	ps	**y**
omega	Ω	*mathematical constant*	ω	*angular frequency*	long o	**w**

stigma	ς					**v**
pomega			ϖ	*angular velocity*		**v**

rg

Counting

No.	***Greek***	***Latin***
1	*mono*	*uni*
2	*duo*	*bi*
3	*tri*	*tri*
4	*tetra*	*quad*
5	*penta*	*quin*
6	*hexa*	*sex*
7	*hepta*	*sept*
8	*octo*	*oct*
9	*nona*	*non*
10	*deca*	*dec*

These booklets are written and produced by Robert Goodhand

Although the formulae and expressions given have been individually derived and checked errors do creep in. The booklets are also continuously updated.

If you would like the latest issue, just email me at robert.goodhand@gmail.com