1) Regression, correlation and hypothesis testing

1.1) Exponential models

1.2) Measuring correlation

1.3) Hypothesis testing for zero correlation

1.1) Exponential models

Chapter CONTENTS

Wor	ked e	exan	nple		Your turn									
The table shows some in °C, of a colony of ba	The table shows some data collected on the temperature, in °C, of a colony of bacteria (t) and its growth rate (g) .													
Temperature, t 3 (°C)	5	6	8	9	11	Temperature, t (°C)	3	5	6	8	9	11		
Growth rate, <i>g</i> 1.40	1.94	1.97	2.85	3.2	4.64	Growth rate, g	1.04	1.49	1.79	2.58	3.1	4.46		
The data are coded us and $y = \log g$. The reg y = -0.0536 + 0.063 a) Find the initial grow b) Given that the data the form $g = kb^t$ w values of k and b .	ing the c ression 7 <i>x</i> . vth rate can be r vhere <i>k</i> a	hanges line of y modelle and <i>b</i> ai	of vari	able <i>x</i> = s found	= <i>t</i> to be ion of ind the	The data are code and $y = \log g$. Th y = -0.2215 + 0 a) Find the initial b) Given that the the form $g = x$ values of k and a) 0.6 b) $k = 0.6, b =$	ed usir e regr 0.0792 grow data o kb^t w d b. 1.20	ng the oression <i>x</i> . th rate can be here <i>k</i>	thanges line of f	s of vari y on x is ed by ar re cons	able <i>x</i> = s found n equat tants, f	= <i>t</i> to be ion of ind the		

Worked example	Your turn
A rabbit population, <i>P</i> , is modelled with respect to time in years, <i>t</i> . An exponential model is proposed: $P = kb^t$ The data is coded using $x = t$ and $y = \log P$. The regression line of <i>y</i> on <i>x</i> is found to be $y = 3 + 0.2x$. Determine the values of <i>k</i> and <i>b</i> .	A rabbit population, <i>P</i> , is modelled with respect to time in years, <i>t</i> . An exponential model is proposed: $P = kb^t$ The data is coded using $x = t$ and $y = \log P$. The regression line of <i>y</i> on <i>x</i> is found to be $y = 2 + 0.3x$. Determine the values of <i>k</i> and <i>b</i> . k = 100, b = 2.00 (3 sf)

1.2) Measuring correlation

Chapter CONTENTS

Worked example	Your turn								
Calculate the product moment	Calculate the product moment								
correlation coefficient for the following	correlation coefficient for the following								
data:	data:								
x y 1 3 2 4 3 5 4 8									

Worked example								Your turn														
From the large data set, the daily mean temperature, $t °C$, and the daily total rainfall, $r mm$, were recorded from 27^{th} May to 5^{th} June inclusive 1987 in Leuchars.								Fi ki re	From the large data set, the daily mean windspeed, <i>w</i> knots, and the daily maximum gust, <i>g</i> knots, were recorded for the first 10 days in September in Hurn in 1987.													
Day 1 2 3 4 5 6 7 8 9 10										Day	1	2	3	4	5	6	7	8	9	10		
t	8. 5	9. 0	10.3	12.8	13.5	12.8	9. 8	8.8	10.0	10.4		w	4	4	8	7	12	12	3	4	7	10
r	0	2.4	8.1	0.2	0.4	tr	6.1	3.6	tr	31.8		g	13	12	19	23	33	37	10	n/a	n/a	23
a) S b) C fr c) V t d	tate t alcula or the r' rea Vith ro he su ata.	the m the ten ading efere itabil	he produced days, s.	ng of oduct , stati	tr in t t mon ing cle ur ans ear re	the tanent of early b swer t	ible a corre how to pa tion r	above Platio you o rt b, node	e. n coe deal v comn el for t	efficient with the nent on these	a b c a b c) S () C () M ()	tate alcul or the vith r ne su ata. or the or the c is c or re nd c ata neal	the m ate th e rem refere litabil on d nese 0.953 lose latio daily n point r regr	ainin ainin ity of aily i days 33 (2 to 1 n be max ts lie ressi	ng of oduc g 8 c to yo a lin max f sf) so t so t twe imur clos	n/a II t mor lays. ur an: ear re imun imun here en da n gu se to node	n the ment swer egres: n gus is st aily n st. Th a str l is s	to pa sion r st is r rong nean nis m raigh uitat	abov lation rt b, c node not a s pos win ieans t line ole.	ve. comn l for t vaila itive dspe s tha e, so	fficient nent on hese ble ed t the a

1.3) Hypothesis testing for zero correlation Chapter CONTENTS

Worked example	Your turn
A scientist takes 19 observations of the masses of two reactants in an experiment. She calculates a product moment correlation coefficient of $r = 0.54$.	A scientist takes 14 observations of the masses of two reactants in an experiment. She calculates a product moment correlation coefficient of $r = -0.45$.
The scientist believes there is a positive correlation between the masses of the two reactants. Test at the 1% level of significance, the scientist's claim, stating your hypotheses clearly.	The scientist believes there is a negative correlation between the masses of the two reactants. Test at the 5% level of significance, the scientist's claim, stating your hypotheses clearly. $H_0: \rho = 0$ $H_1: \rho < 0 \therefore$ One-tailed test. Sample size = 14 Significance level in tail = 5% Reject H_0 if $r < -0.4575$ r = -0.45 > -0.4575 The result is not significant. Insufficient evidence to reject H_0 . Insufficient evidence to suggest there is a negative correlation between the masses of the two reactants.

Worked example	Your turn
A scientist takes 20 observations of the masses of	A scientist takes 30 observations of the masses of
two reactants in an experiment.	two reactants in an experiment.
She calculates a product moment correlation	She calculates a product moment correlation
coefficient of $r = 0.54$.	coefficient of $r = -0.45$.
The scientist believes there is no correlation	The scientist believes there is no correlation
between the masses of the two reactants.	between the masses of the two reactants.
Test at the 1% level of significance, the scientist's	Test at the 10% level of significance, the scientist's
claim, stating your hypotheses clearly.	claim, stating your hypotheses clearly.
	$H_0: \rho = 0$ $H_1: \rho \neq 0$: Two-tailed test. Sample size = 30 Significance level in each tail = 5% Reject H_0 if $r < -0.3061$ r = -0.45 < -0.3061 The result is significant. Sufficient evidence to reject H_0 . Sufficient evidence to suggest there is a correlation between the masses of the two reactants.

Worked example									Your turn											
The table from the large data set shows the daily mean temperature, $t {}^{\circ}C$, and the daily total rainfall, $r mm$, in Leuchars for a sample of nine days in October 1987.							The table from the large data set shows the daily maximum gust, x knots, and the daily maximum relative humidity, y %, in Leeming for a sample of eight days in May 2015.													
t	11.4	10.5	6.5	8.3	8.2	5.7	7.6	12.1	11.2		x	31	28	38	37	18	17	21	29	
Test, at the 10% level of significance, whether there is evidence of a negative correlation between daily mean temperature and daily total rainfall. State your hypotheses clearly.									y9994878080898486Test, at the 10% level of significance, whether there is evidence of a positive correlation between daily maximum gust and daily maximum relative humidity. State your hypotheses clearly. $H_0: \rho = 0$ $H_1: \rho > 0 \therefore$ One-tailed test. Sample size = 8											
										Reject H r = 0.11 The results Insufficities Correlation	Ince I ₀ if 149 ult is ent ent ion im r	r > < 0 s not evid evid betv elati	0.50 .506 sigr lence lence veer ve h	each)67 nifica e to r e to s n dail umic	int. rejec sugg y ma lity.	= 10 t H ₀ est t axim	here um g	e is a (ust a	posi and d	tive daily