Version MH/6

EGT2
ENGINEERING TRIPOS PART IIA

Thursday 2 May 2024 14.00 to 15.40

Module 3E3

MODELLING RISK

*Answer not more than **two** questions.*

All questions carry the same number of marks.

*The **approximate** percentage of marks allocated to each part of a question is indicated in the right margin.*

*Write your candidate number **not** your name on the cover sheet.*

STATIONERY REQUIREMENTS
Single-sided script paper

SPECIAL REQUIREMENTS TO BE SUPPLIED FOR THIS EXAM
CUED approved calculator allowed.
Attachment: 3E3 Modelling Risk data sheet (5 pages).

10 minutes reading time is allowed for this paper at the start of the exam.

You may not start to read the questions printed on the subsequent pages of this question paper until instructed to do so.

You may not remove any stationery from the Examination Room.

1 (a) BeCool, one of the most popular cafés in Trapani, has a single drive-thru lane. Customers arrive at the lane, place their orders, wait for their drinks to be prepared, and then pick them up and leave. BeCool management team wants to understand how long customers wait on average and how many baristas they need to schedule to provide good service. It is known that customers arrive with an average of 10 customers per hour (0.167 per minute). The coefficient of variation of arrival times is 1. There is 1 barista working, who takes an average of 3 minutes to prepare a drink. Customers cannot leave the queue until their drinks are ready.

(i) What is the queuing system description of BeCool? Explain it in terms of arrivals, departures, and queues. [10%]

(ii) How many baristas should BeCool schedule to ensure that average waiting time is no more than 2 minutes? [10%]

(iii) BeCool management team decides to add a separate lane for priority customers (e.g., delivery services) employing 1 skilled barista. Priority customers arrive at the same rate as regular customers, but their drinks take only 2 minutes to prepare on average. The coefficient of variation of arrival times of priority customers is also 1. Regular customers still experience the same preparation time as before. What is the expected waiting time in queue for regular and priority customers under the new system? [20%]

(b) The historical monthly sales data (January to June) for a small retail store is 50, 60, 55, 70, 65, 80.

(i) Using the 3-month simple moving average, forecast the sales for July. Show your calculations. [10%]

(ii) Using exponential smoothing, forecast the sales for July. Use a smoothing factor of $\alpha = 0.3$. Show your calculations. Interpret the smoothing factor. [20%]

(iii) In time series analysis and forecasting, define key terms including trend components and cyclical patterns. Illustrate your understanding by presenting a practical example. [10%]

(c) Describe a Markov chain with a practical example. [20%]

2 (a) Answer the following two questions in a context of portfolio management.

(i) Draw a typical case in a risk return-chart. Use the chart to explain the classification of investment options depending on the risk. Explain the factor that influence your preferred strategy to invest. [15%]

(ii) Let's suppose now the case of two statistically independent shares in the portfolio, X and Y, each costing £10/share. The distributions of returns are $X \sim N\ (\mu = 12,\ \sigma^2 = 9)$ and $Y \sim N\ (\mu = 11,\ \sigma^2 = 16)$, hence the portfolio return is also normally distributed. What is the probability of 50-50 portfolio losing 25% or more of its value? Compare it to a 100% investment in Y. Which is safer? Why? [15%]

(b) As the founder of a tech start-up you invested £100,000 in developing a prototype for a new software product. To bring the product to market, an additional investment of £150,000 is needed for marketing, testing, and user feedback implementation. If the product is launched, it is successful (i.e. 'favourable condition') with probability 0.6, bringing a revenue of £400,000. However, it can also be unsuccessful (i.e. 'unfavourable condition'), in which case the revenue will only be £50,000. You still can choose not to launch the product to the market and recover £75,000.

(i) Develop a decision tree that represents your options and potential outcomes. Assess the financial feasibility of moving forward with the product launch. [10%]

(ii) Analyse how changes in the probability of positive market response would impact your decision to proceed or abandon the project. Specify the threshold probability at which your decision would change. [10%]

(iii) Determine the threshold cost at which you would reconsider launching the product. [10%]

(iv) Given potential uncertainties in the market, an external market research firm provides additional insights from the national software assessment department. The department provides evaluations of 'desirable' and 'undesirable' conditions, conditional to the market status they find, 'favourable' and 'unfavourable', with conditional probabilities: P('desirable' | 'favourable conditions') = 0.8; and P('undesirable' | 'unfavourable conditions') = 0.9. Given a 'desirable' assessment under 'favourable conditions', what is the maximum amount you should be willing to pay for acquiring additional market research insights? [20%]

(c) Identify and describe three main pitfalls encountered in the field of data analytics. Provide examples or explanations for each pitfall. [20%]

 (TURN OVER

3 (a) Consider a scenario where a marketing analyst is studying the relationship between advertising expenditure (X) and sales (Y) for a retail company. The analyst collects data from different regions and performs a linear regression analysis. The following ANOVA table (Table 1) summarises the results.

Table 1

Source	**Sum of Squares**	**Deg. of freedom**	**Mean Square**	**F-Statistic**	**p-value**
Regression	120,000	1	√	√	$p < 0.05$
Residual	80,000	18	√		
Total	√	√			

(i) Fill in the blanks, marked by √, in the ANOVA table. State the null hypothesis for the F-test in the context of this regression analysis. Given the degrees of freedom for regression and residual, determine the sample size used in the analysis. [10%]

(ii) Calculate the coefficient of determination (R^2) for this regression analysis. Interpret its meaning in the context of the study. List the main steps of a hypothesis test for the effect of advertising on sales, including the null and alternative hypotheses and interpret the p-value. [20%]

(iii) Provide a 95% confidence interval for the slope (regression coefficient) in the model, assuming the estimated slope coefficient from your model is 20. [10%]

(b) Consider the case of an engineering student, Taylor, who runs a small 3D printing service on their university campus. Taylor plans to optimise the production of 3D-printed prototypes, at a production cost of £10 per unit, to maximise profits while minimising the risk of excess inventory. He decides to sell each prototype at £15 unit. The probability distribution of the demand for 3D-printed prototypes is a normal distribution of mean 50 and standard deviation 6. Taylor, who knows about the newsvendor problem, targets an order quantity initially of $Q = 55$.

(i) Taylor assumes that any unsold 3D-printed prototypes can be sold to the biological sciences laboratory after the production cycle. If the biological sciences laboratory buys entire leftover prototypes at £v per prototype, where v is less than the production cost, what is the value of v? [20%]

(ii) Suppose the salvage value for unsold prototypes is such that the underage and overage costs are equal. What would you recommend to Taylor as optimal order quantity? Provide a brief explanation for your answer. [10%]

 (cont.

(iii) Taylor learns that there is an upcoming engineering project with tight deadlines, likely increasing the demand for 3D-printed prototypes. Based on this information, Taylor comes up with the discrete demand distribution of Table 2. What is the optimal order quantity under this new demand forecast, considering that all unsold prototypes can be salvaged at the engineering laboratory for £7 per prototype?

Table 2

Q	50	60	70	80	90	100
$P(X = Q)$	0.15	0.20	0.25	0.18	0.12	0.10

[10%]

(c) Define the term "bottleneck" in the context of process analysis. How can bottlenecks negatively impact a business? How can they be identified? What steps should be taken to identify their root causes? [20%]

END OF PAPER

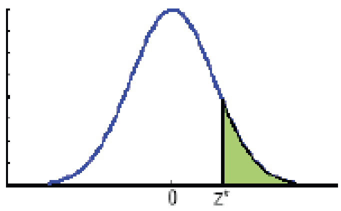
THIS PAGE IS BLANK

Version MH/6

EGT2
ENGINEERING TRIPOS PART IIA
2 May 2024, Module 3E3, Questions 1-3.

SPECIAL DATA SHEET

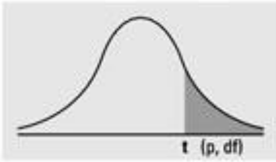
Standard Normal distribution table: *Areas under the standard normal curve beyond z*, i.e., shaded area.*



z*	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010

t-distribution table:

Numbers in each row of the table are values on a *t*-distribution with (*df*) degrees of freedom for selected right-tail (greater-than) probabilities (*p*).



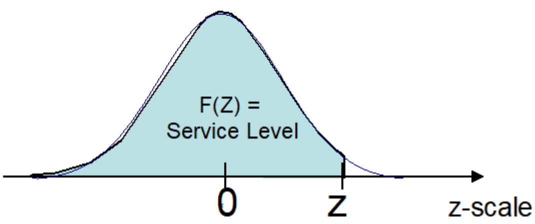
df/p	0.40	0.25	0.10	0.05	0.025	0.01	0.005	0.0005
1	0.324920	1.000000	3.077684	6.313752	12.70620	31.82052	63.65674	636.6192
2	0.288675	0.816497	1.885618	2.919986	4.30265	6.96456	9.92484	31.5991
3	0.276671	0.764892	1.637744	2.353363	3.18245	4.54070	5.84091	12.9240
4	0.270722	0.740697	1.533206	2.131847	2.77645	3.74695	4.60409	8.6103
5	0.267181	0.726687	1.475884	2.015048	2.57058	3.36493	4.03214	6.8688
6	0.264835	0.717558	1.439756	1.943180	2.44691	3.14267	3.70743	5.9588
7	0.263167	0.711142	1.414924	1.894579	2.36462	2.99795	3.49948	5.4079
8	0.261921	0.706387	1.396815	1.859548	2.30600	2.89646	3.35539	5.0413
9	0.260955	0.702722	1.383029	1.833113	2.26216	2.82144	3.24984	4.7809
10	0.260185	0.699812	1.372184	1.812461	2.22814	2.76377	3.16927	4.5869
11	0.259556	0.697445	1.363430	1.795885	2.20099	2.71808	3.10581	4.4370
12	0.259033	0.695483	1.356217	1.782288	2.17881	2.68100	3.05454	43178
13	0.258591	0.693829	1.350171	1.770933	2.16037	2.65031	3.01228	4.2208
14	0.258213	0.692417	1.345030	1.761310	2.14479	2.62449	2.97684	4.1405
15	0.257885	0.691197	1.340606	1.753050	2.13145	2.60248	2.94671	4.0728
16	0.257599	0.690132	1.336757	1.745884	2.11991	2.58349	2.92078	4.0150
17	0.257347	0.689195	1.333379	1.739607	2.10982	2.56693	2.89823	3.9651
18	0.257123	0.688364	1.330391	1.734064	2.10092	2.55238	2.87844	3.9216
19	0.256923	0.687621	1.327728	1.729133	2.09302	2.53948	2.86093	3.8834
20	0.256743	0.686954	1.325341	1.724718	2.08596	2.52798	2.84534	3.8495
21	0.256580	0.686352	1.323188	1.720743	2.07961	2.51765	2.83136	3.8193
22	0.256432	0.685805	1.321237	1.717144	2.07387	2.50832	2.81876	3.7921
23	0.256297	0.685306	1.319460	1.713872	2.06866	2.49987	2.80734	3.7676
24	0.256173	0.684850	1.317836	1.710882	2.06390	2.49216	2.79694	3.7454
25	0.256060	0.684430	1.316345	1.708141	2.05954	2.48511	2.78744	3.7251
26	0.255955	0.684043	1.314972	1.705618	2.05553	2.47863	2.77871	3.7066
27	0.255858	0.683685	1.313703	1.703288	2.05183	2.47266	2.77068	3.6896
28	0.255768	0.683353	1.312527	1.701131	2.04841	2.46714	2.76326	3.6739
29	0.255684	0.683044	1.311434	1.699127	2.04523	2.46202	2.75639	3.6594
30	0.255605	0.682756	1.310415	1.697261	2.04227	2.45726	2.75000	3.6460
z	0.253347	0.674490	1.281552	1.644854	1.95996	2.32635	2.57583	3.2905
CI	———	———	80%	90%	95%	98%	99%	99.9%

Z-Chart & Loss Function

F(Z) is the probability that a variable from a standard normal distribution will be less than or equal to Z, or alternately, the service level for a quantity ordered with a z-value of Z.

L(Z) is the standard loss function, i.e. the expected number of lost sales as a fraction of the standard deviation. Hence, the lost sales = L(Z) x σ_{DEMAND}

Z	F(Z)	L(Z)	Z	F(Z)	L(Z)	Z	F(Z)	L(Z)	Z	F(Z)	L(Z)
-3.00	0.0013	3.000	-1.48	0.0694	1.511	0.04	0.5160	0.379	1.56	0.9406	0.026
-2.96	0.0015	2.960	-1.44	0.0749	1.474	0.08	0.5319	0.360	1.60	0.9452	0.023
-2.92	0.0018	2.921	-1.40	0.0808	1.437	0.12	0.5478	0.342	1.64	0.9495	0.021
-2.88	0.0020	2.881	-1.36	0.0869	1.400	0.16	0.5636	0.324	1.68	0.9535	0.019
-2.84	0.0023	2.841	-1.32	0.0934	1.364	0.20	0.5793	0.307	1.72	0.9573	0.017
-2.80	0.0026	2.801	-1.28	0.1003	1.327	0.24	0.5948	0.290	1.76	0.9608	0.016
-2.76	0.0029	2.761	-1.24	0.1075	1.292	0.28	0.6103	0.274	1.80	0.9641	0.014
-2.72	0.0033	2.721	-1.20	0.1151	1.256	0.32	0.6255	0.259	1.84	0.9671	0.013
-2.68	0.0037	2.681	-1.16	0.1230	1.221	0.36	0.6406	0.245	1.88	0.9699	0.012
-2.64	0.0041	2.641	-1.12	0.1314	1.186	0.40	0.6554	0.230	1.92	0.9726	0.010
-2.60	0.0047	2.601	-1.08	0.1401	1.151	0.44	0.6700	0.217	1.96	0.9750	0.009
-2.56	0.0052	2.562	-1.04	0.1492	1.117	0.48	0.6844	0.204	2.00	0.9772	0.008
-2.52	0.0059	2.522	-1.00	0.1587	1.083	0.52	0.6985	0.192	2.04	0.9793	0.008
-2.48	0.0066	2.482	-0.96	0.1685	1.050	0.56	0.7123	0.180	2.08	0.9812	0.007
-2.44	0.0073	2.442	-0.92	0.1788	1.017	0.60	0.7257	0.169	2.12	0.9830	0.006
-2.40	0.0082	2.403	-0.88	0.1894	0.984	0.64	0.7389	0.158	2.16	0.9846	0.005
-2.36	0.0091	2.363	-0.84	0.2005	0.952	0.68	0.7517	0.148	2.20	0.9861	0.005
-2.32	0.0102	2.323	-0.80	0.2119	0.920	0.72	0.7642	0.138	2.24	0.9875	0.004
-2.28	0.0113	2.284	-0.76	0.2236	0.889	0.76	0.7764	0.129	2.28	0.9887	0.004
-2.24	0.0125	2.244	-0.72	0.2358	0.858	0.80	0.7881	0.120	2.32	0.9898	0.003
-2.20	0.0139	2.205	-0.68	0.2483	0.828	0.84	0.7995	0.112	2.36	0.9909	0.003
-2.16	0.0154	2.165	-0.64	0.2611	0.798	0.88	0.8106	0.104	2.40	0.9918	0.003
-2.12	0.0170	2.126	-0.60	0.2743	0.769	0.92	0.8212	0.097	2.44	0.9927	0.002
-2.08	0.0188	2.087	-0.56	0.2877	0.740	0.96	0.8315	0.090	2.48	0.9934	0.002
-2.04	0.0207	2.048	-0.52	0.3015	0.712	1.00	0.8413	0.083	2.52	0.9941	0.002
-2.00	0.0228	2.008	-0.48	0.3156	0.684	1.04	0.8508	0.077	2.56	0.9948	0.002
-1.96	0.0250	1.969	-0.44	0.3300	0.657	1.08	0.8599	0.071	2.60	0.9953	0.001
-1.92	0.0274	1.930	-0.40	0.3446	0.630	1.12	0.8686	0.066	2.64	0.9959	0.001
-1.88	0.0301	1.892	-0.36	0.3594	0.605	1.16	0.8770	0.061	2.68	0.9963	0.001
-1.84	0.0329	1.853	-0.32	0.3745	0.579	1.20	0.8849	0.056	2.72	0.9967	0.001
-1.80	0.0359	1.814	-0.28	0.3897	0.554	1.24	0.8925	0.052	2.76	0.9971	0.001
-1.76	0.0392	1.776	-0.24	0.4052	0.530	1.28	0.8997	0.047	2.80	0.9974	0.001
-1.72	0.0427	1.737	-0.20	0.4207	0.507	1.32	0.9066	0.044	2.84	0.9977	0.001
-1.68	0.0465	1.699	-0.16	0.4364	0.484	1.36	0.9131	0.040	2.88	0.9980	0.001
-1.64	0.0505	1.661	-0.12	0.4522	0.462	1.40	0.9192	0.037	2.92	0.9982	0.001
-1.60	0.0548	1.623	-0.08	0.4681	0.440	1.44	0.9251	0.034	2.96	0.9985	0.000
-1.56	0.0594	1.586	-0.04	0.4840	0.419	1.48	0.9306	0.031	3.00	0.9987	0.000
-1.52	0.0643	1.548	0.00	0.5000	0.399	1.52	0.9357	0.028			



Z & *L(z)* for special service levels

Service Level F(z)	*z*	*L(z)*
75%	0.67	0.150
90%	1.28	0.047
95%	1.64	0.021
99%	2.33	0.003

Inventory management:

(Q,R) model - optimal solution:

$$Q = \sqrt{\frac{2\lambda[K+pn(R)]}{h}}; \qquad F(R) = 1 - \frac{Qh}{p\lambda}$$

Newsvendor model:

$$F(Q^*) = \frac{c_u}{c_u + c_o}$$

Queuing theory:

Values of L_q for s servers, with mean utilisation rate ρ, assuming Poisson arrivals and exponential service times (known as a *M/M/s* queue).

Utilisation rate	***Number of servers (s)***				
(ρ)*	**1**	**2**	**3**	**4**	**5**
.10	.0111	.0020	.0004	.0001	.0000
.20	.0500	.0167	.0062	.0024	.0010
.30	.1286	.0593	.0300	.0159	.0086
.35	.1885	.0977	.0552	.0325	.0196
.40	.2667	.1524	.0941	.0605	.0398
.45	.3682	.2285	.1522	.1052	.0743
.50	.5000	.3333	.2368	.1739	.1304
.55	.6722	.4771	.3583	.2772	.2185
.60	.9000	.6750	.5321	.4306	.3542
.62	1.0116	.7743	.6213	.5109	.4269
.64	1.1378	.8880	.7246	.6051	.5130
.66	1.2812	1.0188	.8446	.7158	.6152
.68	1.4450	1.1698	.9847	.8461	.7367
.70	1.6333	1.3451	1.1488	1.0002	.8816
.72	1.8514	1.5500	1.3423	1.1834	1.0553
.74	2.1062	1.7914	1.5721	1.4025	1.2646
.76	2.4067	2.0785	1.8472	1.6668	1.5187
.78	2.7655	2.4237	2.1803	1.9887	1.8302
.80	3.2000	2.8444	2.5888	2.3857	2.2165
.82	3.7356	3.3661	3.0979	2.8832	2.7029
.84	4.4100	4.0265	3.7456	3.5190	3.3273
.86	5.2829	4.8852	4.5914	4.3526	4.1493
.88	6.4533	6.0414	5.7345	5.3834	5.2682
.90	8.1000	7.6737	7.3535	7.0898	6.8624
.92	10.5800	10.1392	9.8056	9.5290	9.2893
.94	14.7267	14.2712	13.9240	13.6344	13.3821
.96	23.0400	22.5698	22.2088	21.9060	21.6408
.98	48.0200	47.5350	47.1602	46.8439	46.5656
.99	98.0101	97.5176	97.1357	96.8127	96.4274

Multi-server waiting in queue

$$W_q = \frac{1}{\mu s} \frac{\rho^{\sqrt{2(s+1)}-1}}{1-\rho} \frac{CV_a^2 + CV_s^2}{2}$$

Variance of a portfolio:

Consider three random variables x, y and z with means m_x, m_y, and m_z, respectively; variances Var(x), Var(y), and Var(z), respectively; and covariance between x and y, for example, given by the formula above. Given any numbers α_x, α_y, α_z, let $v = \alpha_x x + \alpha_y y + \alpha_z z$. Then the variance of v is given by

$$\mathrm{Var}(v) = \alpha_x{}^2\mathrm{Var}(x) + \alpha_y{}^2\mathrm{Var}(y) + \alpha_z{}^2\mathrm{Var}(z) + 2\left(\alpha_x\alpha_y\mathrm{cov}(x,y) + \alpha_y\alpha_z\mathrm{cov}(y,z) + \alpha_x\alpha_z\mathrm{cov}(x,z)\right)$$

Covariance, correlation, and regression:

Consider data pairs $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$.

Let m_X and m_Y denote the respective means of the X and Y data.

Let s_X and s_Y denote the respective standard deviations of the X and Y data.

Covariance between X and Y is given by

$$\mathrm{cov}(X,Y) = \frac{\sum_{i=1}^{n}(X_i - m_X)(Y_i - m_Y)}{n} = \frac{\sum_{i=1}^{n} X_i Y_i}{n} - m_X m_Y$$

The correlation coefficient between X and Y is given by

$$\mathrm{correl}(X,Y) = r = \frac{\mathrm{cov}(X,Y)}{s_X s_Y}.$$

The line of best fit is given by

$$Y - m_Y = \frac{r s_Y}{s_X}(X - m_X).$$