

# Answers to Exercises in Chapter 4 - Statistics and Curve Fitting

- 4-1. The `rand` function in MATLAB produces uniformly-distributed pseudo-random numbers. The command `rand(20,1)` produces this sequence of numbers.

$$\left\{ \begin{array}{l} 0.9501, 0.2311, 0.6068, 0.4860, 0.8913, 0.7621, 0.4565, 0.0185, 0.8214, 0.4447 \\ 0.6154, 0.7919, 0.9218, 0.7382, 0.1763, 0.4057, 0.9355, 0.9169, 0.4103, 0.8936 \end{array} \right\}$$

- (a) Compute the sample mean. 0.6237
- (b) Estimate the standard deviation of the numbers produced by the `rand` function from this set of 20 numbers. 0.2824
- (c) Estimate the standard deviation of the sample mean from the answers in (a) and (b). 0.0632
- (d) Approximately how many data are required to make an estimate of the mean with a standard deviation less than 0.02? 706
- 4-2.  $N$  independent samples are obtained from a Gaussian-distributed random variable and used to calculate a sample variance  $S_x^2$  which is an estimate of  $\sigma_x^2$ , the actual variance of the random variable.  $S_x^2$  is a random variable and therefore has an expected value  $E(S_x^2)$  and a variance  $\text{Var}(S_x^2)$ . How large should  $N$  be in order to make the standard deviation of  $S_x^2$  less than one percent of the actual variance  $\sigma_x^2$  of the random variable? >2002
- 4-3. Repeat Exercise 4-2 with everything the same except the samples are obtained from a uniform distribution. >8002
- 4-4. A manufacturer of photodiodes takes data over a long period of time and establishes that the responsivity of the photodiodes has a standard deviation of 5% and that the responsivity distribution is Gaussian.
- (a) If a customer wants to establish the average responsivity of photodiodes of this type to within  $\pm 1\%$  with 95% confidence how many photodiodes must he test? >96
- (b) If the customer tests 11 photodiodes, what would be his confidence limit on the average responsivity of the photodiodes, based on these measurements alone, if he wants to be 95% confident?  $\pm 3.3\%$ .
- 4-5. In the population of men in the United States the mean body weight  $E(w)$  is 80 kg. A researcher is testing men in San Diego to determine whether or not their mean body weight is the same as in the general population. The null and alternate hypotheses are  $H_0: E(w) = 80 \text{ kg}$ ,  $H_1: E(w) \neq 80 \text{ kg}$  and the acceptance region is  $78 \text{ kg} \leq \bar{w} \leq 82 \text{ kg}$  where  $\bar{w}$  is the sample mean.
- (a) If a sample of 20 men is used and the actual mean body weight of men in San Diego is 80 kg, the standard deviation is 15 kg and the distribution is Gaussian, what is the probability of a type I error? 0.551
- (b) If a sample of 200 men is used and the actual mean body weight of men in San Diego is 80 kg, the standard deviation is 15 kg and the distribution is Gaussian, what is the probability of a type I error? 0.0594
- (c) If a sample of 200 men is used and the actual mean body weight of men in San Diego is 77 kg, the standard deviation is 12 kg and the distribution is Gaussian, what is the probability of a type II error? 0.1193
- 4-6. Using the data below find fits of  $Y$  to  $X$  with polynomials of degrees 1 through 8 and decide which degree is the best to use based on the variation of mean-squared error with degree.
- | $X$    | $Y$     |
|--------|---------|
| 0.5548 | -1.0074 |
| 0.1210 | 0.1070  |

0.4508	-0.8648
0.7159	-0.7739
0.8928	-0.5361
0.2731	-0.3784
0.2548	-0.2685
0.8656	-0.7230
0.2324	-0.4466
0.8049	-0.6901
0.9084	-0.4477
0.2319	-0.2316
0.2393	-0.2500
0.0498	0.7378
0.0784	0.6172
0.6408	-0.9550
0.1909	0.0015
0.8439	-0.7546
0.1739	-0.0378
0.1708	-0.0362

5<sup>th</sup> degree is probably the best fit

- 4-7. An integrator is tested to determine its transfer function. Measurements are made at multiple frequencies and the measurements have random noise on them. The measurements of transfer function magnitude versus radian frequency are

$\omega$	$ H(j\omega) $
10.0000	1.8982
12.9155	1.7066
16.6810	1.2068
21.5443	0.9965
27.8256	0.8212
35.9381	0.6799
46.4159	0.4020
59.9484	0.3766
77.4264	0.2527
100.0000	0.2368

It is known that the transfer function of an integrator is of the form  $H(j\omega) = K / j\omega$  where  $K$  is a constant. If the magnitude of the transfer function is expressed in dB it is

$$|H(j\omega)|_{\text{dB}} = 20\log_{10}(K/\omega) = 20\log_{10}(K) - 20\log_{10}(\omega) = K_{\text{dB}} - \omega_{\text{dB}}.$$

This equation is linear in  $20\log_{10}(\omega)$ . If we let  $X = \omega_{\text{dB}}$  and  $Y = |H(j\omega)|_{\text{dB}}$  then the linearized model for the transfer function magnitude is

$$Y = K_{\text{dB}} - X.$$

Assuming that  $\omega$  is known exactly, then  $X$  is also known exactly and the only unknown is  $K_{\text{dB}}$ . Given the noisy data, find the best estimate of  $K_{\text{dB}}$ , by minimizing the mean-squared error between  $Y$  and  $K_{\text{dB}} - X$ . 21.3529

- 4-8. For diodes of a certain doping profile their junction capacitance  $C$  varies with applied reverse-bias voltage  $V$  according to

$$C = \sqrt{\frac{K_1}{K_2 + V}}.$$

which can be rearranged into  $VC^2 = K_1 - K_2C^2$ . Measurements are made of capacitance and reverse-bias voltage and the measurements have random errors.

V	C (pF)
1.0000	12.3737
2.0000	9.3408
3.0000	8.0870
4.0000	7.5832
5.0000	6.6589
6.0000	6.2304
7.0000	5.9242
8.0000	4.9676
9.0000	4.9742
10.0000	4.7156

Using linear curve fitting, find the best estimates of  $K_1$  and  $K_2$ .

$$K_1 = 241.915 \text{ and } K_2 = 0.5831.$$

- 4-9. In calibrating a pressure sensor it is discovered that its calibration is a function of its temperature. To compensate for the temperature sensitivity a thermocouple is added to the sensor so that the sensor output consists of two signals, a voltage from the pressure sensor and a voltage from the thermocouple. Then the new combined sensor is calibrated by subjecting it to several combinations of known pressures and temperatures. The calibration data are summarized below. Find a calibration equation in which the pressure sensor voltage,  $V_p$ , and the thermocouple voltage,  $V_T$ , are the independent variables and pressure,  $P$ , is the dependent variable. Let the highest degree of pressure voltage in the calibration equation be, "np" and let the highest degree of thermocouple voltage be "nt". Make a table of the mean-squared difference between the calibration pressures and the calculated pressures from the calibration equation versus "nt" and "np" for all combinations of "nt" and "np" from 0 through 3 (16 total combinations, some are given below).

#### Calibration Data

Pressure	Thermocouple Voltage	Pressure Sensor Voltage
0	0.78506	0.1066
5	0.76672	1.6687
10	0.81147	2.795
15	0.72724	4.7024
20	0.93708	7.1971
0	1.0806	0.62316
5	0.99712	2.2758
10	1.1052	3.842
15	1.1903	6.1822
20	1.2069	7.7139
0	1.2369	1.5871
5	1.1929	3.2299
10	1.3791	5.3615
15	1.4573	7.5834
20	1.3652	8.9165
0	1.4323	2.288
5	1.5796	5.1485
10	1.6297	6.9672
15	1.6339	8.3048
20	1.5668	10.032

	nt = 0	nt = 1	nt = 2	nt = 3
np = 0	50.0000	48.2231	48.0476	47.9948
np = 1	9.7504	0.7528	0.4339	0.4330
np = 2	9.6468	0.7034	0.4154	0.4077
np = 3	9.6229	0.7023	0.4153	0.4071