Semester 1 2017,

Example Exam 1

Solved

Instructions

The exam consists of 4 questions, 1-4. Each question has four items, a-d.

Within each question:

Item (a) carries a weight of 8 marks.

Item (b) carries a weight of 7 marks.

Item (c) carries a weight of 6 marks.

Item (d) carries a weight of 4 marks.

The total marks in the exam are 100.

Answer ALL questions in the spaces provided.

If more space is required, use the back of the PREVIOUS page.

Show all your working and include sketches where appropriate.

Work written in the formulae and tables booklet will NOT be marked.

Question 1: $\boxed{}$ from formula sheet

The time it takes a computer system to boot is an exponentially distributed random variable X with variance 4.

(a) Determine the rate parameter, λ and use it to write an expression for the cumulative distribution function of X.

For exponential,

$$\boxed{\sigma^2 = \frac{1}{\lambda^2}} = 4, \quad \lambda^2 = \frac{1}{4} \Rightarrow \lambda = 1/2$$

$$\boxed{\overline{F(x)} = 1 - F(x) = e^{-\lambda x}}$$

$$F(x) = 1 - e^{-\lambda x}$$
$$= 1 - e^{-x/2}$$

(b) Assume that operating costs during system boot are X^2 (the square of the boot time). What is the expected value of the operating costs?

want $E(X^2)$

$$\boxed{\mathrm{var}(X) = E(X^2) - [E(X)]^2} \Rightarrow E(X^2) = \mathrm{var}(X) + [E(X)]^2$$
$$= 4 + 4$$
$$= 8$$

$\mathrm{var}(X) = 4$

$\boxed{E(X) = 1/\lambda} = 2$

$[E(X)]^2 = 4$

(c) Assume now that 10 such independent systems are booting in parallel. Let N denote the number of systems that have booted before time 2. What is the probability that $N > 1$?

let $p = pr$(system boots before 2)

$$pr(X \leq 2) = 1 - e^{-2/2} = 1 - e^{-1} = 0.632$$

$$N \sim bin(10, 0.632)$$

$$pr(N > 1) = 1 - pr(N \leq 1)$$

$$= 1 - [pr(N=0) + pr(N=1)]$$

$$= 1 - (1-0.632)^{10} - 10(1-0.632)^{9} \cdot 0.632^{1}$$

$$= 0.999$$

(d) Describe in words how you would write Monte-Carlo simulation code that estimates your answer to (c).

- simulate 10 exponential random variables
- count how many of these are less than 2. call this N.

repeat these steps 10^6 times, count how many N are > 1, use this to calculate probability. This should be close to value in (c).

Question 2:

The weights of the left and right wing of an aircraft, denoted (X, Y), are distributed as a bivariate Normal random variable. It is known that $E(X) = E(Y) = 1000$ (kilograms) and $\text{var}(X) = \text{var}(Y) = 9$. Further the correlation coefficient is $\rho = 0.9$.

(a) Find the covariance between X and Y.

$$\boxed{\text{cov}(X,Y) = \rho_{xy}\,\sigma_x\,\sigma_y}$$

$$= 0.9 \times \sqrt{9} \times \sqrt{9}$$

$$= 8.1$$

(b) Let $T = X + Y$ be the total weight of the wings. Find an expression in terms of the standard Normal cdf, $\Phi(\cdot)$, for the probability that $T > t$ for any t.

$$\boxed{E(X+Y) = E(X) + E(Y)} = 1000 + 1000 = 2000$$

$$\boxed{\text{Var}(X+Y) = \text{var}(X) + \text{var}(Y) + 2\,\text{cov}(X,Y)} = 9 + 9 + 2\times 8.1 = 34.2$$

$$T \sim N(2000, 34.2)$$

$$\Pr(T > t) = \Pr\left(Z > \frac{t-2000}{\sqrt{34.2}}\right) = 1 - \Pr\left(Z \leq \frac{t-2000}{\sqrt{34.2}}\right)$$

$$= 1 - \Phi\left(\frac{t-2000}{\sqrt{34.2}}\right)$$

(c) Assume now that the weight of the non-wing parts of the aircraft, denoted U, is Normally distributed with mean 3000 and variance 16. Thus the total aircraft weight is $W = U + X + Y$. Assume further that U is independent of X and that U is independent of Y. It turns out that flight costs are proportional to $0.001 * W^2$. Calculate the expected flight costs.

$E(W) = E(U) + E(X+Y) = 3000 + 2000 = 5000$

$var(W) = var(U) + var(X+Y) = 16 + 34.2 = 50.2$

want $E(0.001 W^2) = 0.001 E(W^2)$

$E(W^2) = var(W) + [E(W)]^2$

$= 50.2 + 5000^2$

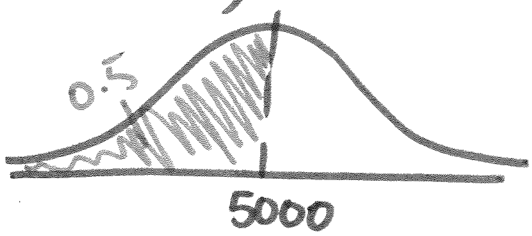
$= 25\,000\,050.2$

$E(0.001 W^2) = 25\,000.0502$

(d) 10 aircraft are manufactured yielding an i.i.d. sequence of total weights, $W_1, \ldots, W_{10}$, each distributed as W from (c). An aircraft is considered light if $W < 5000$. Let N denote the number of light aircraft. What is the variance of N?

$W \sim N(5000, 50.2)$

$N \sim bin(10, 0.5)$ $\quad pr(W < 5000) = 0.5$



$\boxed{var(N) = np(1-p)}$

$= 10 \times 0.5 \times 0.5$

$= 2.5$

Question 3:

You are investigating the erosion rate on two types of building materials. You obtain data for the erosion rate of material type X from 8 samples and material type Y from 10 samples. The sample means are 8 and 5 and the sample standard deviations are 2.3 and 2.1 (all units in mm/year).

(a) Assume that you have reason to believe that the population variances are the same. Calculate the pooled standard deviation.

$$\boxed{S_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}}$$

$$= \sqrt{\frac{(8-1)\times 2.3^2 + (10-1)\times 2.1^2}{18-2}}$$

$$= \sqrt{4.795}$$

$$= 2.19$$

(b) Construct a 99% confidence interval for the difference in mean wear between the two building materials.

$$\boxed{\bar{x} - \bar{y} \pm t_{n_1+n_2-2;1-\alpha/2}\, S_p \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$$t_{16;0.995} = 2.921$$

$$= 2.3 - 2.1 \pm 2.921 \times 2.19 \times \sqrt{\frac{1}{8}+\frac{1}{10}}$$

$$= 0.2 \pm 3.034$$

$$= (-2.834, 3.234)$$

(c) You wish to test whether the mean wear is the same for both materials. Carry out an Hypothesis test with $H_0 : \mu_X = \mu_Y$ vs. $H_1 : \mu_X \neq \mu_Y$. Assume $\alpha = 10\%$ and state your conclusion.

$$\boxed{ts = \frac{\bar{x} - \bar{y} - 0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}} = \frac{0.2}{2.19 \times \sqrt{\frac{1}{8} + \frac{1}{10}}}$$

$$= 0.192$$

0.192

From tables $1.337 > 0.19 > 0$

$2 \times 0.1 < p < 2 \times 0.5$

p-value $> \alpha = 0.1$, $\therefore$ retain H_0.

There is no significant evidence of a difference in mean wear.

(d) Suggest (in words or code) how to write code for evaluating the power of the test (e.g. using Monte Carlo Simulation), assuming that $\mu_X = 7.5$, $\mu_Y = 5.5$ and $\sigma^2 = 4$ for both populations.

simulate data:

sample1 = rand(Normal(7.5,4), 8)
sample2 = rand(Normal(5.5,4), 10)
calc test stat (ts)

repeat this 10^6 times and compare to the critical value. count how many simulations give |ts| larger than crit val to compute probability. This is the power of the test.

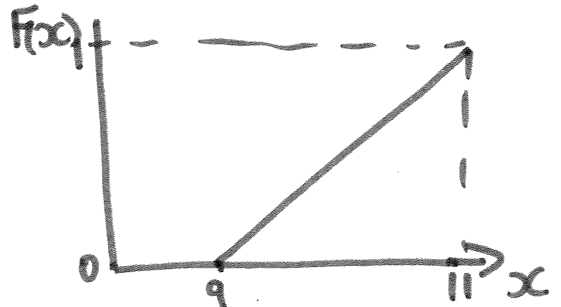
Question 4:

The height of a brick is uniformly distributed with a mean of 10 cm and a variance of 1/3.

(a) Sketch the cdf. $H \sim unif(a,b).$

Find $a,b.$

$$\left.\begin{aligned} \boxed{E(X) = \frac{a+b}{2}} = 10 &\Rightarrow a+b=20 \\ \boxed{var(X) = \frac{(b-a)^2}{12}} = \frac{1}{3} &\Rightarrow (b-a)^2 = 4 \\ &\quad (b-a) = 2 \end{aligned}\right\} \Rightarrow \begin{aligned} b &= 11 \\ a &= 9 \end{aligned}$$



(b) Five independent bricks are placed above each other. What is the variance of the total height?

$$\begin{aligned} var(X_1 + X_2 + X_3 + X_4 + X_5) &= var(X_1) + \dots + var(X_5) \\ &= 1/3 + \dots + 1/3 \\ &= 5/3 \end{aligned}$$

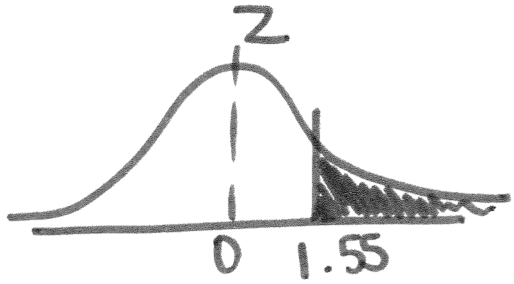
(c) Use the central limit theorem to approximate the probability that the total height of the bricks is greater than 52 cm.

let T be total height

$E(T) = 10+10+10+10+10 = 50$

$T \sim_{approx} N(50, 5/3)$

$Pr(T>52) = Pr\left(Z > \frac{52-50}{\sqrt{5/3}}\right)$



$= Pr(Z > 1.55) = 1 - Pr(Z \leq 1.55)$

$= 1 - 0.9394$

$= 0.0606$

(d) 5 such bricks are randomly selected. What is the probability that 4 out of the 5 bricks are higher than 10.5cm (that is, one is less than 10.5cm)?

let N be # bricks > 10.5 $\quad Pr(H > 10.5) = 0.25$

$N \sim bin(5, 0.25)$

$pr(N=4) = {}^5C_4 \times 0.25^4 \times 0.75^1$

$= 0.0146$