MIDDLE EAST TECHNICAL UNIVERSITY
FACULTY OF ENGINEERING

CE 204 UNCERTAINTY and DATA ANALYSIS
Spring Semester 2013-14

Homework 3- Date Due: April 28th, 2014 Monday till 16.00

IMPORTANT NOTICE:

- *You are allowed to collaborate with other students (or ask questions to your assistants/ instructors) on homework provided that you stay away from plagiarizing (according to dictionaries "to plagiarize" means to steal and pass off ideas and/or words/ solutions of another as one's own without citing the source). That is, collaboration is accepted if you write and give your own solutions. If you are caught on plagiarizing or cheating by handing in "too similar" homework, you will be graded by zero on this homework.*
- *For late submission of homework, there will be 10 points of deduction and no homework will be accepted after Apil, 29th 16.00 pm.*

1. The total depth of snow fall over a ski resort during winter months is modelled using log-normal distribution with mean and standard deviation as 38cm and 8cm, respectively.

a) Find the parameters of the log-normal distribution.

b) Past experience tells without sufficient natural snow cover over the resort (winter snow > 28cm), the management has to rely on expensive artificial snow making techniques. What is the probability in any given year that artificial snow will be necessary?

c) On the other hand extended snow storms (winter snow > 56) over the resort is also a problem as it prevents ski lovers to reach the resort. On average what is the return period of such storms over the resort?

SOLUTION

a) $\zeta^2 = \ln(1+\delta^2) = \ln\left(1+\left(8/38\right)^2\right) = 0.0434 \Rightarrow \zeta = 0.2082$

$\lambda = \ln\mu - 1/2\,\zeta^2 = \ln(38) - 1/2 \times 0.0434 = 3.6153$

b) $P(x<28) = P\left(z < \frac{\ln 28 - 3.6159}{0.2082}\right) = P(z < -1.362) = 0.0866$

c) $P(x>56) = P\left(z > \frac{\ln 56 - 3.6159}{0.2082}\right) = P(z > 1.966) = 1 - 0.975$

$\cong 0.025$

Return period: $1/0.025 = 40$ years

2. Arrival of a bus at a bus-station during rush-hour traffic is modelled with Poisson process. On the average 4 buses arrive in 20 minutes at this station.

a) What are the chances that any passenger will wait 10 minutes without seeing any bus arriving?

b) There are four different routes using the same bus-station: A, B, C, and D. Given there are twice more buses using route A than B, C, and D during rush-hour traffic (i.e. P(A) = 2 P(B) = 2 P(C) = 2 P(D)) and assuming a passenger is planning on catching bus using route A only, what are the chances that this passenger will wait for 20 minutes? (arrival of buses and route selection are statistically independent).

c) During rush hour the buses arriving to this particular bus stop are often completely full, so they cannot accept new passengers. Past experience shows that the probability of the buses to be full is 60%. In this case, passengers keep waiting for the next bus. What are the chances that the passenger waiting for a bus using route B will be able to get on the third bus?

a) For 10 minutes: $\lambda = 2$

$P(x=0) = e^{-\lambda}\lambda^0/0! = e^{-2} \cdot 2^0/1 = 0.1353$

b) For 20 minutes: $\lambda = 4$

$P(A \cap \text{No bus in 20 minutes}) = \underset{0.4}{P(A)} \cdot e^{-4}\,4^0/0! = 0.4 \times 0.0183 = 0.0073$

c) $0.2 \times 0.6 \times 0.6 \times 0.4$

3. The design criteria for a comprehensive harbour construction require tsunami walls of 12 meters high to protect populated coastal areas. Historical records show that there were 8 tsunamis with surge at least 12-meter high during past 89 years. Occurrence of such tsunamis are assumed to obey Poisson process.

a) What is average return period?
b) Find the return period of tsunamis with surge greater than 12-meters.
c) What are the chances of observing one such tsunami within the return period?
d) What are the chances that there will be no such tsunamis during the next 5 years?
e) What are the chances that there will be at least 3 such tsunamis during the next 5 years?
f) What is the difference between Poisson process and Exponential distribution?

SOLUTION

a) The average or mean time between instances of severe/rare events is called the return period.

b) $T = 89/8 \approx 11$ years

c) $\lambda = 89/8 \times 8/89 = 1$; $P(x=1) = \frac{e^{-\lambda} \cdot 1}{1} = 0.3679$

d) $\lambda = 40/89$; $P(x=0) = e^{-40/89} \lambda^0/0! = 0.638$

e) $\lambda = 40/89$; $P(x \geq 3) = 1 - P(x \leq 2) = 1 - \left(0.638 + e^{-40/89}(40/89) + e^{-40/89}(40/89)^2/2!\right) = 1 - 0.9891 = 0.0109$.

f) Poisson $\rightarrow$ Discrete / Exponential $\rightarrow$ Continuous.

4. The inter-arrival times of heavy vehicles at a toll bridge are observed to be random variables obeying exponential distribution with mean 1.5 minutes.

a) What is the probability of observing a heavy vehicle in 30 seconds?
b) If you have waited for 3 minutes and observed no heavy vehicles, what is the probability that a heavy vehicle will be observed in the next 30 seconds? Compare your results in part "a" and "b". State the lack of memory property of exponential distribution in general.

SOLUTION

a) $\theta = 1/1.5_{\text{min}}$ & $f_X(x) = 1/1.5\, e^{-1/1.5 x}$, $x > 0$

$P(x < 30\text{sec}) = P(x < 0.5\text{min}) = \int_0^{0.5} 1/1.5\, e^{-1/1.5 x}\, dx$

$= 0.2835$.

b) $P(3 < x < 3.5 / x > 3) = \frac{P(3 < x < 3.5)}{P(x > 3.0)} = \frac{\int_3^{3.5} 1/1.5\, e^{-1/1.5 x}\, dx}{\int_3^{\infty} 1/1.5\, e^{-1/1.5 x}\, dx}$

$= \frac{0.03836}{0.1353} = 0.2835$.

"a" & "b" give the same result since exponential dist. is memoryless.

$P(T > a+t / T > a) = P(T > t)$

5.The random variable X has a probability density function as

$f_X(x) = c\,x + 0.5 \qquad 1 \le x \le 2$
$f_X(x) = 0 \qquad$ elsewhere.

a) Find the constant c so that $f_X(x)$ is a proper probability density function of the random variable X.
b) If $y = x^2$ find the probability density function of the random variable y and its expected value.

SOLUTION

a) $\int_1^2 (cx+0.5)\,dx = 1 \Rightarrow \quad c x^2/2 + 0.5x \Big|_1^2 = c(2-1/2) + 0.5(2-1) = 1$

$c = 0.5/1.5 = 1/3$.

b) $f_X(x) = 1/3\, x + 0.5, \quad 1 \le x \le 2, \quad F_X(x) = \int_1^x (1/3\, x^* + 0.5)\,dx^*$

$F_X(x) = 1/6\, x^{*2} + 0.5x^* \Big|_1^x = -0.6667 + 0.5x + 1/6\, x^2$
$= -2/3 + 1/2\, x + 1/6\, x^2$.

$y = x^2 \Rightarrow x = \mp\sqrt{y}$, but $1 < x < 2 \Rightarrow x = +\sqrt{y}$

$F_y(y) = -2/3 + 1/2\sqrt{y} + 1/6\, y \qquad , \quad 1 \le y < 4$

$f_y(y) = \frac{dF_y}{dy} = 1/4\, y^{-1/2} + 1/6 = 1/6 + \frac{1}{4\sqrt{y}}, \qquad 1 \le y \le 4$

$= 0$, elsewhere.

5. The moment of inertia of a rectangular section having random sides "a" and "b" with means 10 cm and 20 cm and equal coefficient of variations is given as $I = (a\,b^3)/12$.

a) Find the first order mean value and coefficient of variation of the moment of inertia.
b) Find the second order mean value of the moment of inertia.

SOLUTION

$I \approx \bar{a}\bar{b}^3/12 + \bar{b}^3/12\,(a-\bar{a}) + \frac{3\bar{a}\bar{b}^2}{12}(b-\bar{b}) + 0 + \frac{6\bar{a}\bar{b}}{12\times 2}(b-\bar{b})^2$.

a) First order mean value of I : $\bar{I}_o = \frac{\bar{a}\bar{b}^3}{12} = 6666.67\ cm^4 = \bar{I}_1$

First order variance of I : $V(I)_1 = \bar{b}^6/12^2\, V(a) + \frac{9\bar{a}^2\bar{b}^4}{12^2} V(b)$.

$V(a) = (10\times\delta)^2$ & $V(b) = (20\times\delta)^2$

$V(I)_1 = \frac{20^6}{144}(10\times\delta)^2 + \frac{9\times 10^2\times 20^4}{12^2}(20\times\delta)^2 = 4.444\times 10^8\,\delta^2\,(cm^4)^2$

$\sigma_{I_1} = 21081.9\,\delta\ cm^4$, $\qquad \delta_{I_1} = \frac{21081.9\,\delta}{6666.67} \doteq 3.162\,\delta$.

b) $E(I)_2 = 6666.67 + 0 + 0 + \frac{6\bar{a}\bar{b}}{12\times 2}\underbrace{E(b-\bar{b})^2}_{V(b)}$

$= 6666.67 + \frac{6\times 10\times 20}{24}(20\times\delta)^2$

$= (6666.67 + 20000\,\delta^2)\ cm^4$.

7. There are two sources of electrical power generation. The amount of power generated by each of them are random and are denoted by the variables X and Y. The joint probability density function of X and Y is modelled as

$f_{XY}(x,y) = a(5 - x + y) \qquad 1 < x < 3,\ 0.5 < y < x$
$f_{XY}(x,y) = 0$ elsewhere

a) Find the constant "a" so that $f_{XY}(x,y)$ is a proper joint density function.
b) Find the marginal probability density functions of X and Y.
c) Are X and Y statistically independent? Why yes/why no?
d) Find the correlation coefficient ρ_{XY} and give your comments on its value.
e) Find the conditional probability density function of X given Y (i.e. $f_{X/Y}(x,y)$).
f) What is the conditional expectation of X given Y.
(i.e. $E(x/y) = \int x\, f_{X/Y}(x,y)\,dx$)

SOLUTION

a) $\int_1^3\int_{0.5}^{x} a(5-x+y)\,dy\,dx = 1 \Rightarrow a\int_1^3 (5y - xy + y^2/2)\Big|_{0.5}^{x}\,dx$

$= a\int_1^3 \left(5(x-0.5) - x(x-0.5) + 1/2\,(x^2-0.5^2)\right)dx$

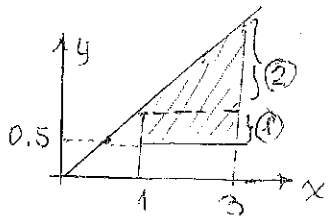
$= a\left(5x^2/2 - 2.5x - x^3/3 + 0.25x^2 + 1/6\,x^3 - \frac{0.25x}{2}\right)\Big|_1^3 = 12.4167a = 1$

$a = 1/12.4167 = 0.0805. \quad (0.0805367)$

$f_{XY}(x,y) = 0.0805(5-x+y), \quad 1<x<3, \quad 0.5<y<x.$
$= 0$, elsewhere.

b) $f_X(x) = \int_{0.5}^{x} 0.0805(5-x+y)\,dy = -0.2114 + 0.443x - 0.0403x^2, \quad 1<x<3.$
$= 0$ elsewhere.

$f_Y(y)_{(1)} = \int_1^3 0.0805(5-x+y)\,dx$



$= 0.4832 + 0.1611y, \quad 0.5<y<1$

$f_Y(y)_{(2)} = \int_y^3 0.0805(5-x+y)\,dx$

$= 0.8456 - 0.1611y - 0.0403y^2, \quad 1<y<3.$
$= 0$ elsewhere

c) $f_{XY}(x,y) \neq f_X(x)\cdot f_Y(y) \Rightarrow \therefore$ X & Y are not statistically independent

d) $COV(XY) = E(XY) - E(X).E(Y)$

$E(XY) = \int_1^3\int_{0.5}^{x} xy\ 0.0805(5-x+y)\,dy\,dx = 3.2496$

$E(X) = \int_1^3 x(-0.2114 + 0.443x - 0.0403x^2)\,dx = 2.1874$

$E(Y) = \int_{0.5}^{1} y(0.4832 + 0.1611y)\,dy + \int_1^3 y(0.8456 - 0.1611y - 0.0403y^2)\,dy$
$= 0.2282 + 1.1812 = 1.4094$, similarly $E(X^2) = 5.0778$.
$E(Y^2) = 0.1787 + 2.1584 = 2.3371 \Rightarrow V(X) = 0.2909,\ V(Y) = 0.3507$

$\rho_{XY} = \frac{(3.2496 - 2.1874 \times 1.4094)}{\sqrt{0.2909 \times 0.3507}} \cong 0.52$

e) $f_{X/Y}(x,y) = \dfrac{f_{XY}(x,y)}{f_Y(y)}$

① $f^{(1)}_{X/Y}(x,y) = \dfrac{0,9127-0,3221y}{0,48322+0,1611y}$, $1<x<3$; $0.5<y<1$

② $f^{(2)}_{X/Y}(x,y) = \dfrac{0,9127-0,3221y}{0,8456-0,1611y-0,0403y^2}$, $1<x<3$; $1<y<3$.

f) $E(X/Y) = \int_1^3 x \cdot f_{X/Y}(x,y)\,dx$

$= \int_1^3 x\, f^1_{X/Y}\,dx + \int_1^3 f^2_{X/Y}\,dx$.

8. 100 test results in measuring the relations between the random variables X and Y are given in the table below.

Y \ X	0	1	2
0	30	0	30
1	0	40	0

a) Find the joint probability mass function of the random variables X and Y using the table above.
b) Find the correlation coefficient ρ_{XY} and give your comments on its value ?
c) Are X and Y statistically independent? Why yes/why no?
d) Find the conditional expectation of Y given X (i.e. $\sum y\, P_{Y/X}(x, y)$).

SOLUTION

a) $P_{XY}(x,y)$

Y \ X	0	1	2
0	0.3	0	0.3
1	0	0.4	0

$P_X(x=0) = 0.3$, $P_X(x=1) = 0.4$, $P_X(x=2) = 0.3$

$P_Y(y=0) = 0.6$, $P_Y(y=1) = 0.4$

b) $\mathrm{cov}(XY) = E(XY) - E(X).E(Y)$

$E(XY) = 0\times 0.3 + 0 + 0\times 0.3 + 0 + 1\times 1\times 0.4 + 1\times 2\times 0$
$= 0.4$

$E(X) = 0\times 0.3 + 1\times 0.4 + 2\times 0.3 = 1.0$

$E(Y) = 0\times 0.6 + 1\times 0.4 = 0.4.$

$\mathrm{cov}(XY) = 0.4 - 1\times 0.4 = 0$

$\rho_{XY} = 0$

c) $P_{XY}(x=0, y=0) = 0.3 \neq P_X(x=0) \cdot P_Y(y=0)$
So X & Y are not statistically independent.

d) $E(Y/X) = \sum y\, P_{Y/X}(x,y)$

$P_{Y/X}(x,y) = \dfrac{P_{XY}(x,y)}{P_X(x)}$

$P_{Y=0/X=0} = \dfrac{0.3}{0.3} = 1$, $P_{Y=0/X=1} = 0$, $P_{Y=0/X=2} = 1$

$P_{Y=1/X=0} = 0$, $P_{Y=1/X=1} = \dfrac{0.4}{0.4} = 1$, $P_{Y=1/X=2} = 0$

$E(Y/X) = 0\times 1 + 0\times 1 + 1\times 0 + 1\times 1 = 1.$

9. A cantilever beam of length $L = 4$ m supports a point load S_1 applied at the midspan and another point load S_2 applied at its tip as hown in the following figure. The loads are random with means $\mu_1 = \mu_2 = \mu$ and standard deviations, $\sigma_1 = \sigma_2 = \sigma$. The loads are perfectly orrelated (i.e. the correlation coefficient $\rho_{12} = 1.0$). Let the bending moment and the shear force at the fixed end of the cantilever eam be denoted by **M** and **V**, respectively (note that $\mathbf{M} = 2\,S_1 + 4\,S_2$ and $\mathbf{V} = S_1 + S_2$).

a) Find the expected value and variance of **M** and **V**.
b) Determine the value of the correlation coefficient between **M** and **V**. Comment on your result.

SOLUTION

a) $E(M) = 2E(S_1) + 4E(S_2) = 2\mu + 4\mu = 6\mu$

$E(V) = E(S_1) + E(S_2) = 2\mu$

$V(M) = 4V(S_1) + 16V(S_2) + 2\times2\times4\times1.0\times\sigma\times\sigma = 36\sigma^2$

$V(V) = V(S_1) + V(S_2) + 2\times1.0\times\sigma\times\sigma = 4\sigma^2$

b) $COV(MV) / \sigma_M \sigma_V = \rho_{MV}$.

$COV(MV) = E(MV) - E(M)E(V)$

$MV = 2S_1^2 + 2S_1S_2 + 4S_2S_1 + 4S_2^2 = 2S_1^2 + 6S_1S_2 + 4S_2^2$

$E(MV) = 2E(S_1^2) + 6E(S_1S_2) + 4E(S_2^2)$

$E(S_1^2) = V(S_1) + \overline{E(S_1)}^2 = \sigma^2 + \mu^2, \quad E(S_2^2) = V(S_2) + E(S_2)^2 = \sigma^2 + \mu^2$

$E(S_1S_2) - E(S_1)E(S_2) = COV(S_1S_2)$

$\rho_{12} \Rightarrow \underbrace{1.0\times\sigma\times\sigma}_{COV(S_1S_2)} = E(S_1S_2) - \mu^2 \Rightarrow E(S_1S_2) = \sigma^2 + \mu^2$

$E(MV) = 2(\sigma^2+\mu^2) + 6(\sigma^2+\mu^2) + 4(\sigma^2+\mu^2) = 12(\sigma^2+\mu^2)$

$COV(MV) = 12(\sigma^2+\mu^2) - 6\mu\times2\mu = 12\sigma^2$

$$\rho_{MV} = \frac{12\sigma^2}{2\sigma\times6\sigma} = 1.0$$

10. A friction pile is driven through three soil layers and bearing capacity C of the pile is given by:

$$C = 2\, f_{sand} + 3\, f_{silt} + 2\, f_{clay}$$

where f_{sand}, f_{silt} and f_{clay} are random friction forces applied to the pile by each soil layer. All of the friction forces can be modeled as normal variables with the parameters as in the following table.

Variable	Mean (kPa) (μ)	Standard Deviation (kPa) (σ)	Coefficient of Variation (c.o.v.) (δ)
f_{sand}	35	10	
f_{silt}	30		0.15
f_{clay}	45	9	

Please note that the random variable f_{sand} is independent from both random variables f_{silt} and f_{clay}, whereas f_{silt} and f_{clay} are correlated with correlation coefficient, $\rho = 0.25$.

a) Calculate the mean value and standard deviation of the pile bearing capacity, C.
b) Calculate the probability that the pile bearing capacity C is larger than 230 kPa.
c) Calculate the mean value of the load P that the pile can support with 5% probability of failure(i.e. $\Pr(P \geq C) = 0.05$) assuming P as a normal variable with standard deviation $\sigma_P = 3$ kPa.

SOLUTION

a) $E(C) = 2\,E(f_{sand}) + 3\,E(f_{silt}) + 2\,E(f_{clay})$

$= 2 \times 35 + 3 \times 30 + 2 \times 45 = 250$ kPa,

$V(C) = 4\,V(f_{sand}) + 9\,V(f_{silt}) + 4\,V(f_{clay}) + 2 \times 3 \times 2\, \rho_{silt-clay}\, \sigma_{silt}\, \sigma_{clay}$

$= 4(10^2) + 9(30 \times 0.15)^2 + 4(9^2) + 2 \times 3 \times 2 \times 0.25\,(30 \times 0.15) \times 9$

$= 1027.75\ (kPa)^2$

$\sigma_C = 32.059$ kPa,

b) $P(C > 230) = P\left(Z > \dfrac{230 - 250}{32.059}\right) = P(Z > -0.6239)$

$= 1 - 0.2664 = 0.7336$.

c) $P(P \geq C) = 0.05 \Rightarrow P(P - C \geq 0) = 0.05$

$P\left(\dfrac{P - C - E(P-C)}{\sigma_{P-C}} \geq \dfrac{0 - E(P-C)}{\sigma_{P-C}}\right) = 0.05$

$E(P-C) = E(P) - E(C) = E(P) - 250$ kPa

$V(P-C) = V(P) + V(C) = 3^2 + 1027.75 = 1036.75\ (kPa)^2$

$P(P \geq C) = P\left(Z \geq \dfrac{0 - E(P) + 250}{\sqrt{1036.75}}\right) = 0.05$

$\dfrac{250 - E(P)}{32.2} = 1.65 \Rightarrow E(P) = 250 - 1.65 \times 32.2$

$\cong 197$ kPa