2009 Midterm Exam Solution

Problem #1

(1) We should assume a distribution when we want to calculate a confidence interval with the Chevyshev inequality: **F**

(2) The hypergeometricu distribution is based on a Bernoulli process: **F**

(3) The unbiased estimator is always efficient: **F**

(4) The type II error always increases as the type I error decreases: **T**

Problem #2

To reach Grenoble, France, from Turin, Italy, one can follow either of two routes. The first directly connects Turin and Grenoble, whereas the second passes through Chambery, France. During extreme weather conditions in winter, travel between Turin and Grenoble is not always possible because some parts of the highway may not be open to traffic. Denote with A, B, and C the events that the highways from Turin to Grenoble, Turin to Chambery, and Chambery to Grenoble are open, respectively. In anticipation of driving from Turin to Grenoble, a traveler listens to the next day's weather forecast. If snow is forecast for the next day over the southern Alps, one can assume (on the basis of past records) that Pr[A]= 0.6, Pr[B]= 0.7, Pr[C]= 0.5, and Pr[A | BC]= 0.4.

(1) What is the probability that th traveler will be able to reach Grenoble from Turin?

(Solution)

P[C | B]= P[BC}/P[B]= 0.5 → P[BC]= 0.5×P[B]= 0.5×0.7= 0.35

P[A | BC]= P[ABC}/P[BC]= 0.4 → P[ABC]= 0.4×P[B]= 0.4×0.35= 0.14

P[Turin to Grenoble]= P[A] + (P[BC] - P[ABC])= 0.6 + (0.35 - 0.14)= 0.81

(2) What is the probability that traveler will be able to drive from Turin to Grenoble by way of Chambery?

(Solution)

P[BC]= P[C | B]×P[B]= 0.5×0.7= 0.35

(3) Which route should be taken in order to maximize the change of reaching Grenoble?

(Solution)

P[A]= 0.6, P[BC]= 0.35

Selecting route A is better than others

Problem #3

A simple model of annual maximum flood events is $F= 1.2A - 0.6R$, wherein A is the area of the watershed and R is a measure of the runoff. The random variables A and R are both normal, such that $A \sim N[1600, (10)]$ and $R \sim N[2.5, (1.0)]$. Also, Correlation[A, R]= 0.70. What size flood should an engineer design for such that it will be exceeded with a probability of 5% (i.e. $P[F \geq f]= 0.05$)?

(Solution)

$E[F]= 1.2 \times 1600 - 0.6 \times 2.5= 1918.5$

$Var[F]= 1.2^2 \times 10 + 2 \times 1.2(-0.6) \times 0.7 \times \sqrt{10} \times \sqrt{1} + 0.6^2 \times 1= 11.52= 3.40^2$

$f_{0.95}= E[F] + 1.96 \times stdev[F]= 1918.5 + 1.96 \times 3.4= 1925.164$

Problem #4

Barges arrive at a lock at an average of 4 each hour.

(1) If the arrival of barges at the lock can be considered to follow a Poisson process, what is the probability that 6 barges will arrive in 2 hours?

(Solution)

$f_X(x; \lambda)= f_X(6; 8)= 8^6 \times e^{-8}/6!= 0.1221$

(2) If the lock master has just locked through all of the barges at the lock, what is the probability he can take a 15 minute break without another barge arriving?

(Solution)

$f_X(0; 1)= 1^0 \times e^{-1}/0!= 0.3679$

(3) If the operation of the lock is such that 4 barges can be locked through at once and the lock master insist that this always be the case, what is the probability that the first barge to arrive after 4 previous barges have been locked through will have to wait at least 1 hour before being locked through?

(Solution)

$prob(T3 \geq 1)$ is $1-prob(T3 \leq 1)$

The probability that $T \leq 1$ for 3 arrivals comes from the gamma distribution

$P_T(t;\ n,\ \lambda) = \int_0^t p_T(t;\ n,\ \lambda)dt = \int_0^1 (4^3 t^2 e^{-4t}/2!)dt = 0.762$

The desired probability is 1-0.762= 0.238

Problem #5

the strength of asphalt is a concern on a road project. Assume that asphalt has mean strength μ The mean μ can be estimated based on test measurements X_i where $X_i \sim N[\mu, (300)^2]$ Ryan the quality control engineer believes that μ is close to 5,000, so he proposes to employ the average $\bar{X}$ of six observations X_i (i= 1, 2, 3, 4, 5, 6) to construct the estimator

$$\hat{\mu} = (5000 + \bar{X}))/2$$

(1) What are the bias, variance, and mean square error of this estimator as a function of μ?

(Solution)

Bias= $E[\hat{\mu}] - \mu = E[(5000+\bar{X})/2] - \mu = \{(5000 + \mu)/2\} - \mu = (5000 - \mu)/2$

$Var[\hat{\mu}] = Var[(5000+\bar{X})/2] = Var[\bar{X}/2] = (1/4)(\sigma^2/n) = (1/4)(300^2/6) = 3750$

MSE= Var + Bias^2= 3750 + $\{(5000 - \mu)/2\}^2$

(2) Is Ryan's estimator better than just using the sample average $\bar{X}$? Explain.

(Solution)

$Var[\bar{X}] = \sigma^2/n = 15{,}000 > Var[\hat{\mu}]$, and $MSE[\bar{X}] = Var[\bar{X}] = 15{,}000 > MSE[\hat{\mu}]$ if $Bias[\hat{\mu}] < \pm 100$. Therefore, Ryan's estimator has smaller MSE if the true mean really is close to 5,000. Thus, it would make sense to use this estimator if previous tests indicated the t거e value of the mean. On the other hand, if one employed the MVUE criteria. then just using the sample average would be preferred because it is an unbiased estimator of μ