E1 222 Stochastic Models and Applications
Hints for Problem Sheet 3.7

1. Let f be a density function with a parameter θ. (For example, f could be Gaussian with mean θ). Let $X_1, X_2, \cdots, X_n$ be iid with density f. These are said to be an iid sample from f or said to be iid realizations of X which has density f. Any function $T(X_1, \cdots, X_n)$ is called a statistic. Any estimator for θ is such a statistic. We choose a function based on what we think is the best guess for θ based on the sample.

   An estimator $T(X_1, \cdots, X_n)$ is said to be unbiased if $E[T(X_1, \cdots, X_n)] = \theta$. Let us write $\mathbf{X}$ for $(X_1, \cdots, X_n)$ and $T(\mathbf{X})$ for any statistic.

   Suppose θ is the mean of the density f. Show that $T_1(\mathbf{X}) = (X_2 + X_5)/2$, $T_2(\mathbf{X}) = X_1$, $T_3(\mathbf{X}) = (\sum_{i=1}^n X_i)/n$ are all unbiased estimators for θ.

   If T is an estimator for θ, the mean square error of the estimator is $E(T - \theta)^2$. Show that if T is unbiased then the mean square error is equal to the variance of the estimator.

   Among the the three estimators T_1, T_2, T_3 for the mean, listed earlier, which one has least mean square error?

Hint: If T is unbiased, $ET = \theta$ and hence $\mathrm{Var}(T) = E(T - \theta)^2$. We know that average of n iid rv has variance equal to $(1/n)$ times the variance of X_1. Hence, T_3 has least variance and hence least mean-square error (because all of them are unbiased).

2. Let $X_1, \cdots, X_n$ be iid with mean μ and variance σ^2. Let $\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n)$. Show that

$$E\left(\sum_{k=1}^{n} (X_k - \bar{X})^2\right) = (n-1)\sigma^2.$$

   (Hint: Write $(X_k - \bar{X}) = (X_k - \mu) - (\bar{X} - \mu)$ and note that $(\bar{X} - \mu) = \sum_k (X_k - \mu)/n$ and that $E(X_k - \mu)(X_j - \mu) = 0$ for $k \neq j$).

   Based on this, suggest an unbiased estimator for the variance.

   Let $Z = \sum_{k=1}^{n} (X_k - \bar{X})^2$. Suppose the first and third moments of X_i are zero. Find the covariance between $\bar{X}$ and Z.

Hint: The first part involves straight-forward algebra. By using the hint in the problem, you should get it.

We have $\sum_{k=1}^{n} (X_k - \bar{X})^2 = \sum_{k=1}^{n} \left((X_k - \mu) - (\bar{X} - \mu)\right)^2 = \sum_{k=1}^{n} (X_k - \mu)^2 + n(\bar{X} - \mu)^2 - 2(\bar{X} - \mu) \sum_{k=1}^{n} (X_k - \mu)$.

Now, taking expectation on both sides and noting that mean of $\bar{X}$ is μ and its variance is σ^2/n, we get

$$E\left(\sum_{k=1}^{n} (X_k - \bar{X})^2\right) = n\sigma^2 + n\frac{\sigma^2}{n} - 2E\left[(\bar{X} - \mu)\left(\sum_{k=1}^{n} (X_k - \mu)\right)\right]$$

Now, using the hint given in the problem again, we have $(\bar{X} - \mu)\left(\sum_{k=1}^{n} (X_k - \mu)\right) = \frac{1}{n} \sum_{j=1}^{n} (X_j - \mu) \sum_{k=1}^{n} (X_k - \mu) = \frac{1}{n} \left[\sum_{j=1}^{n} \sum_{k \neq j} (X_j - \mu)(X_k - \mu) + \sum_{j=1}^{n} (X_j - \mu)^2\right]$.

Now, taking expectation on both sides and using the hint in the problem,

$$E\left[(\bar{X} - \mu)\left(\sum_{k=1}^{n} (X_k - \mu)\right)\right] = \frac{1}{n}(0 + n\sigma^2)$$

Putting all this together, we get

$$E\left(\sum_{k=1}^{n} (X_k - \bar{X})^2\right) = n\sigma^2 + \sigma^2 - 2\sigma^2 = (n-1)\sigma^2$$

This means that $\frac{1}{n-1} \sum_{k=1}^{n} (X_k - \bar{X})^2$ is an unbiased estimator of variance from n iid samples. This is an important result.

For the second part. Since μ, first moment of X_i, is given to be zero, $E[\bar{X}] = 0$ and hence covariance is $E[\bar{X}Z]$. Multiply the two and then argue that it gives an expression where every term contains either X_j^3 or $X_j^2 X_k$ or $X_i X_j X_k$. Thus expectation is zero because first and third moments are zero and X_i are independent. Thus, $\bar{X}$ and Z are uncorrelated.

3. Let $X_1, X_2, \cdots, X_n$ be iid random variables with mean μ and variance σ^2. Let $\bar{X} = (\sum_{i=1}^{n} X_i)/n$ and $S^2 = \sum_{k=1}^{n} (X_k - \bar{X})^2/(n-1)$ be the sample mean and sample variance respectively. As we have seen, these are unbiased estimators of mean and variance.

   Show that $\mathrm{cov}(\bar{X},\ X_i - \bar{X}) = 0,\ i = 1, 2, \cdots, n$. (Hint: Note that $X_i\bar{X}$ can be written as sum of terms like $X_i X_j$; note that $EX_iX_j = \mu^2$ if

$i \neq j$ and is $\mu^2 + \sigma^2$ if $i = j$; note also that you know mean and variance of $\bar{X}$).

Now suppose that the iid random variables X_i have normal distribution. Show that $\bar{X}$ and S^2 are independent random variables. (Hint: Try to use the result that for jointly Gaussian random variables, uncorrelatedness implies independence).

Hint: Since $E[X_i - \bar{X}] = 0$, we only need to show $E[\bar{X}(X_i - \bar{X})] = 0$. Now, $E[(\bar{X})^2] = \text{Var}(\bar{X}) + (E[\bar{X}])^2 = \frac{\sigma^2}{n} + \mu^2$. We have $X_i\bar{X} = \frac{1}{n}(X_i^2 + \sum_{j \neq i} X_i X_j$. Now it is easy to see that $E[X_i\bar{X}] = \frac{1}{n}(\sigma^2 + \mu^2 + (n-1)\mu^2) = \frac{\sigma^2}{n} + \mu^2$ which shows that $\bar{X}$ and $X_i - \bar{X}$ are uncorrelated.

For the second part, first show that $\bar{X}$ and $X_i - \bar{X}$ are jointly Gaussian for each i. This can be done by writing this as a 2-D vector that can be written as a $2 \times n$ matrix multiplied by the n-dimensional vector with components X_i and noting that X_i are ind and Gaussian and hence jointly Gaussian. Since $\bar{X}$ and $X_i - \bar{X}$ are jointly Gaussian and uncorrelated (from the first part), $\bar{X}$ is independent of $X_i - \bar{X}$ for each i and hence is independent of any function of them.

4. Let X be a nonnegative integer valued random variable. Let $\Phi_X(t) = Et^X$ be its probability generating function and assume that $\Phi_X(t)$ is finite for all t. By arguing as in the proof of Chebeshev inequality, show that for any positive integer, y,

   a. $P[X \leq y] \leq \frac{\Phi_X(t)}{t^y}$, $0 \leq t \leq 1$;

   b. $P[X \geq y] \leq \frac{\Phi_X(t)}{t^y}$, $t \geq 1$.

   Now suppose X is a Poisson random variable with parameter λ. Use the above to show that

$$P\left[X \leq \frac{\lambda}{2}\right] \leq \left(\frac{2}{e}\right)^{\lambda/2}.$$

Hint: You have derived one of the inequalities in the first mid-term test. The second inequality can be proved in a similar form and was shown in answers to mid-term test. The RHS of inequality holds for all t in a range. So, you can minimize the RHS (using the Φ_X of Poisson distribution) over the appropriate range of t to solve the second part of the problem.

5. Let X, Y be two random variables each having mean zero and variance one. Let ρ be the correlation coefficient of X, Y. Show that

$$E[\max(X^2, Y^2)] \leq 1 + \sqrt{1 - \rho^2}$$

(Hint: You can use the identity $\max(a, b) = (a + b + |a - b|)/2$).

Hint: Using the relation for the max given, we have

$$E[\max(X^2, Y^2)] = \frac{1}{2}\left(E[X^2] + E[Y^2] + E\left[\left|X^2 - Y^2\right|\right]\right) = \frac{1}{2}+\frac{1}{2}+\frac{1}{2}E\left[|(X-Y)(X+Y)|\right]$$

Noting that $\text{var}(X - Y) = 2(1 - \rho)$ and $\text{var}(X + Y) = 2(1 + \rho)$, we can now complete the problem by making use of Cauchy-Schwarz inequality.

6. Let $X_1, \cdots, X_n$ be iid random variables with a density (or mass) function having a parameter θ. Let $\mathbf{X} = (X_1, \cdots, X_n)$. Let $T(\mathbf{X})$ be a function of $\mathbf{X}$. As mentioned earlier such T is called a statistic. If the conditional distribution of $\mathbf{X}$ given $T(\mathbf{X})$ does not depend on θ then $T(\mathbf{X})$ is called a *sufficient statistic* for θ. Show that $T(\mathbf{X}) = \sum_{i=1}^{n} X_i$ is a sufficient statistic when X_i are poisson with mean θ.

Hint: Let, $S = \sum_{i=1}^{n} X_i$. All you need to show is that the conditional (joint) mass function of $X_1, \cdots, X_n$, conditioned on S, is independent of θ when X_i are iid poisson with mean θ.

We need conditional probability $P[X_i = x_i, i = 1, \cdots, n | S = s]$.

Given any numbers $x_1, \cdots, x_n, s$, we know that $P[X_1 = x_1, \cdots, X_n = x_n, S = s]$ is zero if $x_1 + \cdots + x_n \neq s$ and it is $P[X_1 = x_1, \cdots, X_{n-1} = x_{n-1}, X_n = s - (x_1 + \cdots + x_n)]$ otherwise.

Now, $P[S = s]$ can be written as sum over all $x_1, \cdots, x_{n-1}$ of the probability $P[X_1 = x_1, \cdots, X_{n-1} = x_{n-1}, X_n = s - (x_1 + \cdots + x_n)]$.

Now, writing the expression for the conditional probability, $P[X_i = x_i, i = 1, \cdots, n | S = s]$, using what is said in the previous two paras, and remembering that X_i are independent and they are Poisson with mean θ, you can complete the solution.

7. Suppose you have to play the following game. You are going to be shown N prizes in sequence. At any time you can either accept the one that is being offered or reject it and choose to see the next prize.

Once you reject a prize you cannot go back to it. At any time you are seeing a prize, all the information you have is the relative rank of the prize that you are being offered, with respect to all the ones that have gone by. That is, when you are seeing the third prize, you know how it ranks with respect to the first and second ones that you have already seen and rejected. Consider the following strategy. You fix an integer k between 1 and N. You reject the first k prizes and then accept the first one that you see which is better than all the ones rejected till that point. (If after the first k, in the remaining $N-k$ chances, you never see a prize that is better than all the ones you had rejected till then, then you would end up rejecting all the prizes). Assuming that all possible orderings of the N prizes are equally likely, calculate the probability that this strategy would get you the best prize. Based on this, suggest what is a good value of k to choose.

Hint: This is a solved example in chapter 3 of the book by S. Ross. It is an example of using the conditioning argument. If you are interested, you can look it up in the book.