Avd. Matematisk statistik

KTH Matematik

TENTAMEN I SF2940 SANNOLIKHETSTEORI/EXAM IN SF2940 PROBABILITY THEORY THURSDAY THE 23^{rd} OF OCTOBER 2008 08.00 a.m.–01.00 p.m.

Examinator: Timo Koski, tel. 790 71 34, e-post: timo@math.kth.se

Tillåtna hjälpmedel Means of assistance permitted: Appendix 2 in A.Gut: An Intermediate Course in Probability. Formulas for probability theory SF2940. Pocket calculator.

You should define and explain your notation. Your computations and your line of reasoning should be written down so that they are easy to follow. Numerical values should be given with the precision of two decimal points. You may apply results stated in a part of an exam question to another part of the exam question even if you have not solved the first part. The number of exam questions (Uppgift) is six(6).

Each question gives maximum ten (10) points. 30 points will guarantee a passing result. The grade Fx (the exam can completed by extra examination) for those with 27−29 points.

Solutions to the exam questions will be available at
`http://www.math.kth.se/matstat/gru/sf2940/`
starting from Thursday 23^{rd} of October 2008 at 02.00 p.m..

The exam results will be announced at the latest on Friday the 7^{th} of November on the announcement board of Matematisk statistik at the entry hall of Institutionen för matematik, Lindstedtsvägen 25.

Your exam paper will be retainable at elevexpeditionen during a period of seven weeks after the date of the exam.

AN ADDITION TO THE COLLECTION OF FORMULAS :
In one of the assignments (Uppgift 1-6) below the following integral may turn out to be useful:

$$\int_0^x \ln u \, du = x \cdot \ln x - x, \quad x > 0.$$

A CORRECTION TO THE COLLECTION OF FORMULAS, Section 2.4: The formula for $\mathrm{Var}(X)$ should read

$$\mathrm{Var}(X) = \mathrm{Var}(E(X|Y)) + E(\mathrm{Var}(X|Y)).$$

Lycka till!

- - - - - - - - - - - - - - - - - - - -

Uppgift 1

The random variables $X_1, X_2, \ldots,$ are I.I.D. and non-negative integer-valued. The random variable $N \in \text{Po}(\lambda)$ is independent of $X_1, X_2, \ldots,$. We set

$$S_N = X_1 + X_2 + \ldots + X_N.$$

If we know that $S_N \in \text{Po}(\mu)$, where $0 < \mu < \lambda$, what is the distribution of an X_k?

(10 p)

Uppgift 2

Let $X_1, X_2, \ldots,$ be I.I.D. random variables with $X_i \in U(0,1)$. Set

$$G_n = (X_1 \cdot X_2 \cdots X_n)^{1/n},$$

i.e., G_n is the geometric mean of $X_1, X_2, \ldots$ and X_n. Show that

$$G_n \xrightarrow{P} e^{-1} \quad \text{as } n \to \infty.$$

(10 p)

Uppgift 3

$X_1, X_2, \ldots,$ is a sequence of random variables such that $E(\mid X_i \mid) < \infty$. Let $\mathcal{F}_n$ be the σ-field generated by $X_1, X_2, \ldots, X_n$. Let $X_1, X_2, \ldots,$ be such that for $n \geq 1$

$$E[X_{n+1} \mid \mathcal{F}_n] = aX_n + bX_{n-1}, \quad X_0 = 0.$$

where $a > 0, b > 0, a + b = 1$. Let

$$S_n = \alpha X_n + X_{n-1}.$$

Find a value for α such that $(S_n)_{n \geq 1}$ is a martingale w.r.t. the filtration $(\mathcal{F}_n)_{n \geq 1}$. Explain your steps of solution carefully. (10 p)

Uppgift 4

Let $N_1 = \{N_1(t) \mid t \geq 0\}$ be a Poisson process with parameter p, $0 < p < 1$, and let $N_2 = \{N_2(t) \mid t \geq 0\}$ be a Poisson process with parameter $q = 1 - p > 0$.
N_1 and N_2 are independent.
When T is the time of the occurrence of the first event in N_1, define

$N_2(T) =$ the number of events that have occurred in N_2 when the first event of N_1 occurs.

a) Find the characteristic function of $N_2(T)$. (9 p)

b) What is the distribution of $N_2(T)$? (1 p)

Uppgift 5

$X = \{X(t) \mid -\infty < t < \infty\}$ is a Gaussian stochastic process. Its mean function is $\mu_X(t) = 0$ for all t and its autocorrelation function is

$$E\left(X(t) \cdot X(s)\right) = R(h) = 6e^{-|h|}, \quad h = t - s.$$

a) Please find

$$E\left(X(t) \mid X(s)\right).$$

(2 p)

b) Please find the probability

$$P\left(X(t) > 1 \mid X(t-1) - X(t-2) = 1\right).$$

(8 p)

Uppgift 6

$N = \{N(t) \mid t \geq 0\}$ is a Poisson process with parameter $\lambda > 0$. Consider for $n = 1, 2, \ldots,$

$$X_n = \frac{1}{n} N(n).$$

a) Use Chebysjev's inequality to find an n such that

$$P\left(\mid X_{n^2} - E\left[X_{n^2}\right] \mid > \epsilon\right) \leq 0.01.$$

(2 p)

b) Show using a Borel-Cantelli lemma that

$$X_{n^2} \overset{a.s.}{\to} \lambda, \quad \text{as} \quad n \to \infty.$$

(8 p)

Avd. Matematisk statistik

KTH Matematik

SOLUTIONS TO THE EXAM IN SF2940 PROBABILITY THEORY 08-10-23

Uppgift 1

From the collection of Formulas we know that the p.g.f. $g_{S_N}(t)$ of $S_N \in \text{Po}(\mu)$ is

$$g_{S_N}(t) = e^{\mu(t-1)}.$$

On the other hand, the general formula for the p.g.f. of a random number of non-negative integer valued random variables is

$$g_{S_N}(t) = g_N\left(g_X\left(t\right)\right) = e^{\lambda(g_X(t)-1)},$$

where $g_X\left(t\right)$ is the p.g.f. of the Xs and where we used the fact that $N \in \text{Po}(\lambda)$. Then we have

$$\mu(t-1) = \lambda(g_X\left(t\right) - 1)$$

$$\Leftrightarrow$$

$$g_X\left(t\right) = \frac{\mu}{\lambda}(t-1) + 1 = \frac{\mu}{\lambda}t + 1 - \frac{\mu}{\lambda}$$

Thus the probabilities of X can be obtained as

$$g_X\left(0\right) = 1 - \frac{\mu}{\lambda}, \quad g_X^{'}\left(0\right) = \frac{\mu}{\lambda}$$

and if $\mu < \lambda$, then these are, indeed, probabilities. $d^n/dt^n g_X\left(t\right) = 0$ for all $n \geq 2$ and all t. Thus

$$P\left(X_k = 0\right) = 1 - \frac{\mu}{\lambda}, P\left(X_k = 1\right) = \frac{\mu}{\lambda}, \quad \mu < \lambda,$$

which is a Bernoulli distribution.

$$\text{ANSWER} : \underline{X_k \in \text{Be}\left(\tfrac{\mu}{\lambda}\right), \quad \mu < \lambda}$$

Uppgift 2

We take the natural logarithm of the geometric mean and thus obtain

$$Y_n = \ln G_n = \frac{1}{n}\left(\ln X_1 + \ln X_2 + \ldots + \ln X_n\right).$$

Since X_i are I.I.D., so are $\ln X_i$ I.I.D., and then the weak law of large numbers tells as $n \to \infty$, that

$$Y_n = \frac{1}{n}\left(\ln X_1 + \ln X_2 + \ldots + \ln X_n\right) \xrightarrow{P} E\left[\ln X_1\right].$$

Since $X_1 \in U(0,1)$, we have

$$E[\ln X_1] = \int_0^1 \ln x dx = [x \ln x - x]_0^1 = -1.$$

Thus

$$Y_n \xrightarrow{P} -1 \quad \text{as } n \to \infty.$$

We have with $g(x) = e^x$ that

$$G_n = g(Y_n)$$

Since $g(x) = e^x$ is continuous everywhere (in particular at $x = -1$), theorem 7.7 in chapter VI of A.Gut: An Intermediate Course gives

$$G_n = g(Y_n) \xrightarrow{P} g(-1) = e^{-1},$$

as was to be proved.

Uppgift 3

It is clear that $E(\mid S_n \mid) < \infty$ for any α and that $(S_n)_{n\geq 1}$ is adapted to the increasing sequence of filtrations $(\mathcal{F}_n)_{n\geq 1}$ for any α.
We need thus to find α such that the so called martingale property holds, i.e.,

$$E[S_{n+1} \mid \mathcal{F}_n] = S_n.$$

Let us now use the suggested formula for S_{n+1} to get

$$E[S_{n+1} \mid \mathcal{F}_n] = E[\alpha X_n + X_{n-1} \mid \mathcal{F}_n] = E[\alpha X_{n+1} \mid \mathcal{F}_n] + E[X_n \mid \mathcal{F}_n]$$

$$= \alpha E[X_{n+1} \mid \mathcal{F}_n] + X_n$$

since X_n is measurable w.r.t. $\mathcal{F}_n$ by construction of $\mathcal{F}_n$. Now we use the assumption about the conditional expectation $E[X_{n+1} \mid \mathcal{F}_n]$ to get

$$= \alpha(aX_n + bX_{n-1}) + X_n$$

$$= (\alpha \cdot a + 1) X_n + \alpha \cdot bX_{n-1}.$$

The comparison of the rightmost expression is with $S_n = \alpha X_n + X_{n-1}$. Hence the desired equality $E[S_{n+1} \mid \mathcal{F}_n] = S_n$ is valid if

$$\alpha = (\alpha \cdot a + 1), \quad \alpha \cdot b = 1.$$

The the leftmost equation holds if $\alpha = 1/(1-a)$, and since $a + b = 1$, even the equation $\alpha \cdot b = 1$ is satisfied by the same choice of α as well.

ANSWER: $\underline{\alpha = \frac{1}{1-a}}$.

Uppgift 4

a)By definition of the characteristic function we have

$$\varphi_{N_2(T)}(t) = E\left[e^{itN_2(T)}\right] = E\left[E\left[e^{itN_2(T)} \mid T\right]\right].$$

We know that $T \in \text{Exp}(1/p)$, i.e., T has a density, for the moment denoted by $f_T(u)$, concentrated on the non-negative values. Then

$$E\left[E\left[e^{itN_2(T)} \mid T\right]\right] = \int_0^\infty E\left[e^{itN_2(T)} \mid T = u\right] f_T(u)\, du$$

But if N_1 and N_2 are independent, then N_2 and T are independent. Hence

$$\int_0^\infty E\left[e^{itN_2(T)} \mid T = u\right] f_T(u)\, du = \int_0^\infty E\left[e^{itN_2(u)}\right] f_T(u)\, du =$$

$$= p\int_0^\infty E\left[e^{itN_2(u)}\right] e^{-pu}\, du$$

as $T \in \text{Exp}(1/p)$ so that $f_T(u) = pe^{-pu}$ for $u \geq 0$. Here we know that $N_2(u) \in \text{Po}\left((1-p)u\right)$ and in Appendix 2 in the Collection of Formulas we find

$$E\left[e^{itN_2(u)}\right] = e^{(1-p)u\left(e^{it}-1\right)} = e^{(1-p)uf(t)}$$

where we set $f(t) = e^{it} - 1$ for ease of writing. Thus we have

$$p\int_0^\infty E\left[e^{itN_2(u)}\right] e^{-pu}\, du = p\int_0^\infty e^{(1-p)uf(t)} e^{-pu}\, du$$

$$= p\int_0^\infty e^{-(p-(1-p)f(t))u}\, du =$$

$$= p\int_0^\infty e^{-(p+(1-p)-(1-p)e^{it})u}\, du =$$

$$= p\int_0^\infty e^{-(1-(1-p)e^{it})u}\, du =$$

$$= p\left[-\frac{1}{1-(1-p)e^{it}} e^{-(1-(1-p)e^{it})u}\right]_0^\infty$$

$$= p\left(\frac{1}{1-(1-p)e^{it}}\right) = \frac{p}{1-(1-p)e^{it}}.$$

ANSWER a): $\underline{\varphi_{N_2(T)}(t) = \frac{p}{1-(1-p)e^{it}}.}$

b)The column of characteristic functions in Appendix 2 shows that $N_2(T) \in \text{Ge}(p)$.

ANSWER b): $\underline{N_2(T) \in \text{Ge}(p).}$

Uppgift 5

a) Since X is a Gaussian process, we find $E[X(t) \mid X(s]$ for $t > s$ in the collection of formulas and by the Doob-Dynkin interpretation of conditional expectation. We have

$$E[X(t) \mid X(s] = \mu_{X(t)} + \rho \frac{\sigma_{X(t)}}{\sigma_{X(s)}} \left(X(s) - \mu_{X(s)} \right)$$

Here $\mu_{X(t)} = \mu_{X(s)} = \mu_X(t) = 0$ and

$$\sigma^2_{X(t)} = \sigma^2_{X(s)} = R(0) = 6e^{-|0|} = 6.$$

and since means are zero,

$$\rho = \frac{\text{Cov}[X(t), X(s]}{\sigma_{X(t)} \cdot \sigma_{X(s)}} = \frac{6e^{-|t-s|}}{6}$$

Hence we have obtained

$$\text{ANSWER a): } \underline{E[X(t) \mid X(s] = e^{-|t-s|} \cdot X(s).}$$

b) We write

$$Y = BX$$

where $Y = (Y_1, Y_2)'$, $X = (X(t), X(t-1), X(t-2))'$ and

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix}.$$

Since the mean function is zero, we have that the mean vector of Y is the zero column vector $\mathbf{0}$ and

$$Y \in N(\mathbf{0}, C_Y),$$

where

$$C_Y = BC_X B',$$

and where from the autocorrelation function

$$C_X = 6 \cdot \begin{pmatrix} 1 & e^{-1} & e^{-2} \\ e^{-1} & 1 & e^{-1} \\ e^{-2} & e^{-1} & 1 \end{pmatrix}.$$

This gives

$$C_Y = 6 \cdot \begin{pmatrix} 1 & e^{-1} - e^{-2} \\ e^{-1} - e^{-2} & 2(1 - e^{-1}) \end{pmatrix}.$$

Then

$$P(X(t) > 1 \mid X(t-1) - X(t-2) = 1) = P(Y_1 > 1 \mid Y_2 = 1)$$

We have that the conditional distribution of Y_1 given $Y_2 = 1$ is found as

$$N\left(\mu_1 + \frac{c_{12}}{c_{22}}(1 - \mu_2), c_{11} - \frac{c_{12}^2}{c_{22}} \right).$$

Here $\mu_1 = \mu_2 = 0$, $c_{11} = 6$, $c_{12} = 6\left(e^{-1} - e^{-2}\right)$ and $c_{22} = 12(1 - e^{-1})$, so that

$$Y_1 \mid Y_2 = 1 \in N\left(\frac{6\left(e^{-1} - e^{-2}\right)}{12(1 - e^{-1})}, 6 - \frac{36\left(e^{-1} - e^{-2}\right)^2}{12(1 - e^{-1})}\right) = N\left(e^{-1}/2, 3(2 - e^{-2}(1 - e^{-1})\right).$$

Therefore

$$\frac{Y_1 - e^{-1}/2}{\sqrt{3(2 - e^{-2}(1 - e^{-1}))}} \mid Y_2 = 1 \in N(0, 1).$$

Thus

$$P\left(Y_1 > 1 \mid Y_2 = 1\right) = P\left(\frac{Y_1 - e^{-1}/2}{\sqrt{3(2 - e^{-2}(1 - e^{-1}))}} > \frac{1 - e^{-1}/2}{\sqrt{3(2 - e^{-2}(1 - e^{-1}))}} \mid Y_2 = 1\right)$$

$$= 1 - \Phi\left(\frac{1 - e^{-1}/2}{\sqrt{3(2 - e^{-2}(1 - e^{-1}))}}\right),$$

where $\Phi(x)$ is the cumulative distribution function of the standard normal random variable. This formula is actually a passing solution to the problem. If we strive for the engineering solution with numerical precision we get

$$\frac{1 - e^{-1}/2}{\sqrt{3(2 - e^{-2}(1 - e^{-1}))}} \approx 0.34$$

and

$$1 - \Phi\left(0.34\right) \approx 0.37$$

ANSWER b): $\underline{P\left(X(t) > 1 \mid X(t-1) - X(t-2) = 1\right) \approx 0.37.}$

Uppgift 6

a) When $N = \{N(t) \mid t \geq 0\}$ is a Poisson process with parameter $\lambda > 0$, then $E\left[N(n)\right] = \lambda\, n$ and $\text{Var}\left[N(n)\right] = \lambda\, n$. If

$$X_{n^2} = \frac{1}{n^2} N(n^2),$$

we have

$$E\left[X_{n^2}\right] = \frac{1}{n^2} E\left[N(n^2)\right] = \frac{1}{n^2}\lambda\, n^2 = \lambda,$$

and

$$\text{Var}\left[X_{n^2}\right] = \frac{1}{n^4}\text{Var}\left[N(n^2)\right] = \frac{1}{n^4}\lambda\, n^2 = \frac{1}{n^2}\lambda.$$

Then Chebysjev's inequality gives

$$P\left(\mid X_{n^2} - \lambda \mid > \varepsilon\right) \leq \frac{1}{\varepsilon^2}\text{Var}\left[X_{n^2}\right] = \frac{\lambda}{\varepsilon^2\, n^2}.$$

Thus, if $n > \sqrt{\frac{100\lambda}{\varepsilon^2}}$, then the desired bound is obtained.

ANSWER a): $\underline{n > \sqrt{\frac{100\lambda}{\varepsilon^2}}.}$

b) From case a) Chebysjev's inequality is

$$P\left(\mid X_{n^2} - \lambda \mid > \varepsilon\right) \leq \frac{1}{\varepsilon^2} \mathrm{Var}\left[X_{n^2}\right] = \frac{\lambda}{\varepsilon^2\, n^2}.$$

We have that $\sum_{n=1}^{\infty} \frac{\lambda}{\varepsilon^2\, n^2} = \frac{\lambda}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\lambda \pi^2}{\varepsilon^2 6} < \infty$. Thus

$$\sum_{n=1}^{\infty} P\left(\mid X_{n^2} - \lambda \mid > \varepsilon\right) < \infty$$

and the first of the Borel-Cantelli lemmas entails that

$$P\left(\mid X_{n^2} - \lambda \mid > \varepsilon \quad \text{i.o.}\right) = 0.$$

Since $\varepsilon > 0$ is arbitrary, this shows clearly that

$$X_{n^2} \stackrel{a.s.}{\rightarrow} \lambda, \quad \text{a}s \quad n \rightarrow \infty,$$

as was asserted.