TENTAMEN I SF2940 SANNOLIKHETSTEORI/EXAM IN SF2940 PROBABILITY THEORY MONDAY THE 14^{TH} OF JANUARY 2008 14.00 p.m.–19.00 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 Tuesday 15^{th} of January 2008 at 10.00 a.m..

The exam results will be announced at the latest on Friday the 25^{th} of January 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.

LYCKA TILL!

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

Uppgift 1

A random variable X has a binomial distribution, i.e., $X \in \text{Bin}\,(n,p)$. The random variable Y is also binomially distributed conditionally on X, i.e., $Y \mid X = k \in \text{Bin}\,(k,p)$.
Show by using the probability generating function that $Y \in \text{Bin}(n,p^2)$. (10 p)

Uppgift 2

$X_1, X_2, \ldots,$ are I.I.D. random variables with the density

$$f_X(x) = \begin{cases} e^{-(x-a)} & \text{if } x \geq a \\ 0 & \text{if } x < a. \end{cases}$$

We set $\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i$ and $Y_n := \min(X_1, \ldots, X_n)$.
a) Show that as $n \to \infty$,

$$\overline{X}_n \xrightarrow{P} a + 1.$$

(3 p)

b) Show that as $n \to \infty$,

$$Y_n \xrightarrow{P} a.$$

(7 p)

Uppgift 3

$X_n \in \mathrm{Ge}(p_n)$, $n = 1, 2, \ldots,$.
a) Establish that

$$F_{X_n}(x) = \begin{cases} 1 - (1 - p_n)^{\lfloor x \rfloor + 1}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0. \end{cases}$$

where $\lfloor x \rfloor$ is the integer part of x (i.e., the largest integer smaller than or equal to x). (2 p)

b) Assume that $\lim_{n\to\infty} np_n = \alpha$, where $\alpha > 0$. Show that as $n \to \infty$,

$$\frac{X_n}{n} \xrightarrow{d} \mathrm{Exp}(1/\alpha).$$

(8 p)

Uppgift 4

A random variable X is said to be *Laplace* distributed with parameter $a > 0$, if its density is

$$f_X(x) = \frac{1}{2a} e^{-|x|/a}, \quad -\infty < x < \infty.$$

We write this as $X \in L(a)$.

a) Find the moment generating function of X . (3 p)

b) $X \in \mathrm{Exp(a)}$ and $Y \in \mathrm{Exp(a)}$ are independent. Set $U := X - Y$. What is the distribution of U ? (7 p)

Uppgift 5

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

$$E(X(t) \cdot X(s)) = R(h) = \max(0, 1 - |h|), \quad h = t - s.$$

a) What is the characteristic function of $X(t) - X(t - 0.5)$? (2 p)

b) Please find the probability

$$P(X(t) - X(t - 0.5) > 2).$$

(8 p)

Uppgift 6

Let $N = \{N(t) \mid t \geq 0\}$ be a Poisson process with intensity $\lambda > 0$. We define a sequence of random variables N_k för $k = 0, 1, 2, \ldots$ by sampling at positive integer values of time and subtraction of a function $\psi(k)$ so that

$$N_k = N(k) - \psi(k), k = 1, 2, 3, \ldots, \quad N_0 := 0.$$

How should $\psi(k)$ be chosen so that the discrete time stochastic process $\{N_k\}_{k=1}^{\infty}$ becomes a a martingale with respect to the family of sigma fields $\mathcal{F}_k$ generated by $N_1, N_2, \ldots, N_k$, i.e., $\mathcal{F}_k = \sigma(N_1, N_2, \ldots, N_k)$? (10 p)

Avd. Matematisk statistik

KTH Matematik

SOLUTIONS TO THE EXAM IN SF2940 PROBABILITY THEORY 08-01-14

Uppgift 1

The probability generating function (p.g.f.) of Y is by definition

$$g_Y(t) = E\left[t^Y\right] = E\left[E\left[t^Y \mid X\right]\right]$$

$$= \sum_{k=0}^{n} E\left[t^Y \mid X = k\right] P(X = k)$$

$$= \sum_{k=0}^{n} (1 - p + pt)^k P(X = k),$$

which is the p.g.f. of X evaluated at $1 - p + pt$, i.e.,

$$= g_X(1 - p + pt) = (1 - p + p(1 - p + pt))^n$$

$$= \left(1 - p + p - p^2 + p^2 t\right)^n$$

$$= \left(1 - p^2 + p^2 t\right)^n,$$

which is the p.g.f. of $\mathrm{Bin}(n, p^2)$.

Uppgift 2

a) Since X_i are I.I.D., the weak law of large numbers tells that

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} E[X],$$

where

$$E[X] = \int_a^\infty x e^{-(x-a)} dx.$$

By change of variable $u = x - a$ we get

$$\int_a^\infty x e^{-(x-a)} dx = \int_0^\infty (u + a) e^{-u} du$$

$$= \int_0^\infty u e^{-u} du + a \int_0^\infty e^{-u} du = 1 + a.$$

This establishes the limit as claimed.

b) We need to prove that for every $\epsilon > 0$

$$P(\mid Y_n - a \mid > \epsilon) \to 0, \quad \text{as } n \to \infty.$$

Since $Y_n \geq a$ we have

$$P(\mid Y_n - a \mid > \epsilon) = P(Y_n - a > \epsilon)$$
$$= P(Y_n > a + \epsilon) = \prod_{i=1}^{n} P(X_i > a + \epsilon),$$

since if $Y_n > a + \epsilon$ then for all i we have $X_i > a + \epsilon$, since and X_i are independent. We have

$$P(X_i > a + \epsilon) = 1 - P(X_i \leq a + \epsilon) = 1 - F_{X_i}(a + \epsilon).$$

It holds that

$$F_{X_i}(a + \epsilon) = \int_a^{a+\epsilon} e^{-(t-a)} dt$$
$$= 1 - e^{-\epsilon}.$$

Thus, since X_i are identically distributed,

$$\prod_{i=1}^{n} P(X_i > a + \epsilon) = e^{-n\epsilon} \to 0$$

as $n \to \infty$, as was to be proved.

Uppgift 3

a) $X_n \in \mathrm{Ge}(p_n)$ means that $F_{X_n}(x) = 0$ for $x < 0$. For $x \geq 0$ we have by definition

$$F_{X_n}(x) = \sum_{0 \leq k \leq x} p_X(k) = \sum_{k=0}^{\lfloor x \rfloor} p_X(k) = p_n \sum_{k=0}^{\lfloor x \rfloor} (1 - p_n)^k$$

$$= p_n \frac{1 - (1 - p_n)^{\lfloor x \rfloor + 1}}{1 - (1 - p_n)}$$

where we used the summation formula for a finite geometric series (c.f., Collection of Formulas) , and thus

$$= 1 - (1 - p_n)^{\lfloor x \rfloor + 1},$$

as was to be proved.
b)
We have from part a) of this problem that

$$F_{X_n/n}(x) = P\left(\frac{X_n}{n} \leq x\right) = P(X_n \leq xn) = F_{X_n}(nx)$$

$$= 1 - (1 - p_n)^{\lfloor xn \rfloor + 1}$$
$$= 1 - (1 - p_n)^{nx} (1 - p_n)^{-<nx>+1}$$

where $< nx >= nx - \lfloor xn \rfloor$, and thus $0 \leq < nx > \leq 1$. We have by assumption and by a standard limit, as $n \to \infty$,

$$(1 - p_n)^{nx} = \left(\left(1 - \frac{np_n}{n}\right)^n\right)^x \to e^{-\alpha x}.$$

Also, as $n \to \infty$,

$$(1-p_n)^{-<nx>+1} = \left(1 - \frac{np_n}{n}\right)^{-<nx>+1} \to 1$$

Thus we have shown for $x > 0$ that

$$F_{X_n/n}(x) \to 1 - e^{-\alpha x}.$$

This proves, by an additional observation about the continuity points of the limiting distribution function, the claim as asserted.

Uppgift 4

a) By definition of the moment generating function we have

$$\psi_X(t) = E\left[e^{tX}\right] = \frac{1}{2a}\int_{-\infty}^{\infty} e^{tx} e^{-|x|/a} dx$$

$$= \frac{1}{2a}\left[\int_{-\infty}^{0} e^{tx} e^{x/a} dx + \int_{0}^{\infty} e^{tx} e^{-x/a} dx\right]$$

$$= \frac{1}{2a}\left[\int_{-\infty}^{0} e^{(t+1/a)x} dx + \int_{0}^{\infty} e^{-((1/a-t)x} dx\right]$$

and for $\mid t \mid < \frac{1}{a}$ we get

$$= \frac{1}{2a}\left[\frac{1}{(t+1/a)} + \frac{1}{((1/a-t)}\right] = \frac{1}{1-(at)^2}.$$

ANSWER a): $\underline{\psi_X(t) = \frac{1}{1-(at)^2}, \quad \mid t \mid < \frac{1}{a}.}$

b) By independence we get

$$\psi_{X-Y}(t) = \psi_X(t) \cdot \psi_Y(-t).$$

It follows by computations similar to ones above that

$$\psi_X(t) = \frac{1}{1-at}, \psi_Y(-t)\frac{a}{1-a(-t)} = \frac{1}{1+at}$$

so that

$$\psi_{X-Y}(t) = \frac{1}{1-(at)^2}.$$

ANSWER b): $\underline{X - Y \in L(a).}$

Uppgift 5

Since $X = \{X(t) \mid -\infty < t < \infty\}$ is a Gaussian stochastic process, the random variable

$$Y := X(t) - X(t-0.5)$$

is for all t a Gaussian random variable. We find its mean and variance.

$$E[Y] = E[X(t) - X(t-0.5)] = E[X(t)] - E[X(t-0.5)] = \mu(t) + \mu(t-0.5) = 0,$$

since the mean function $\mu(t)$ is zero. The variance is found as

$$\mathrm{Var}\left[Y\right] = \mathrm{Var}\left[X(t)\right] + \mathrm{Var}\left[X(t-0.5)\right] - 2\mathrm{Cov}\left[X(t), X(t-0.5)\right]$$

and since the mean is zero, this equals

$$= R(0) + R(0) - 2R(0.5) = 1 + 1 - 2 \cdot 0.5 = 1.$$

Thus $Y \in N(0,1)$. By the Collection of Formulas we get the characteristic function.

$$\text{ANSWER a): } \underline{\varphi_{X(t)-X(t-0.5)}(t) = e^{-t^2/2}.}$$

b) Since by the preceding $Y \in N(0,1)$ we have that

$$P\left(X(t) - X(t-0.5) > 2\right) = 1 - P\left(X(t) - X(t-0.5) \le 2\right) = 1 - \Phi(2) \approx 0.0226.$$

where we used the approximation formula for $Q(x)$ in the Collection of Formulas.

$$\text{ANSWER b): } \underline{P\left(X(t) - X(t-0.5) > 2\right) = 0.023.}$$

Uppgift 6

We need to choose the function $\psi(k)$ so that the martingale property

$$E\left[N_{k+1} \mid \mathcal{F}_k\right] = N_k$$

holds. We have that

$$E\left[N_{k+1} \mid \mathcal{F}_k\right] = E\left[N_{k+1} - N_k + N_k \mid \mathcal{F}_k\right]$$
$$= E\left[N(k+1) - N(k) \mid \mathcal{F}_k\right] - \psi(k+1) + \psi(k) + E\left[N_k \mid \mathcal{F}_k\right]$$
$$= E\left[N(k+1) - N(k)\right] - \psi(k+1) + \psi(k) + E\left[N_k \mid \mathcal{F}_k\right]$$

since $N(k+1) - N(k))$ is independent of $\mathcal{F}_k$, as the Poisson process has independent increments. Because N_k is measurable w.r.t. $\mathcal{F}_k$ we have $E\left[N_k \mid \mathcal{F}_k\right] = N_k$ and then

$$E\left[N_{k+1} \mid \mathcal{F}_k\right] = E\left[N(k+1) - N(k)\right] - \psi(k+1) + \psi(k) + N_k$$

$$= \lambda - \psi(k+1) + \psi(k) + N_k,$$

since $N(k+1) - N(k) \in \text{Po}(\lambda)$. Since $N_0 = 0$ and $N(0) = 0$, we have $\psi(0) = 0$. Thus the martingale property holds for all $k = 1, 2, \ldots$, if the following recursion holds

$$\psi(k+1) = \lambda + \psi(k), \quad k = 1, 2, \ldots.$$

By an induction proof we show that

$$\psi(k) = \lambda \cdot k$$

uniquely solves this recursion.

ANSWER a): $\underline{\psi(k) = \lambda \cdot k, \quad k = 0, 1, 2, \ldots.}$