KTH Matematik

Avd. Matematisk statistik

TENTAMEN I SF2940 SANNOLIKHETSTEORI/EXAM IN SF2940 PROBABILITY THEORY, WEDNESDAY OCTOBER 26, 2016, 08.00-13.00.

Examinator : Boualem Djehiche, tel. 08-7907875, email: boualem@kth.se

Tillåtna hjälpmedel/Permitted means of assistance: Appendix 2 in A. Gut: An Intermediate Course in Probability, Formulas for probability theory SF2940, L. Råde & B. Westergren: Mathematics Handbook for Science and Engineering and pocket calculator.

All used notation must be explained and defined. Reasoning and the calculations must be so detailed that they are easy to follow. Each problem yields max 10 p. 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. 25 points will guarantee a passing result.

If you have received 5 bonus points from the home assignments, you may skip Problem 1(a). If you have received 10 bonus points, you may skip the whole Problem 1.

Solutions to the exam questions will be available at http://www.math.kth.se/matstat/gru/sf2940/ starting from Friday 28 October 2016.

Good luck!

Problem 1

(a) The random variables $X_1, X_2, \ldots,$ are non-negative integer-valued and independent and identically distributed (i.i.d.). The random variable $N \in \text{Po}(a),\ a > 0$, is independent of $X_1, X_2, \ldots,$.

Set

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

If we know that $S_N \in \text{Po}(b)$, where $0 < b < a$, show that for each $k = 1, 2, 3, \ldots,$ $X_k \in \text{Be}(\frac{b}{a})$ i.e. Bernoulli-distributed with parameter b/a. (5 p)

(b) The random variables X_1 and X_2 are independent and $N(0,1)$-distributed. Show that the random variables $\frac{X_1}{X_2}$ and $\sqrt{X_1^2 + X_2^2}$ are independent and determine their distributions. (5 p)

(*Hint: Use polar coordinates.*)

Problem 2

$X_1, X_2, \ldots$ is a sequence of independent and identically distributed random variables with mean zero and variance σ^2. The random variable N is independent of the sequence $(X_n)_{n\geq 1}$ and $N \in \text{Po}(\lambda),\ \lambda > 0$. Show that

(a) $\frac{N}{\lambda} \xrightarrow{P} 1$, when $\lambda \to \infty$. (2 p)

(b) $\frac{N-\lambda}{\sqrt{\lambda}} \xrightarrow{d} N(0,1)$, when $\lambda \to \infty$. (2 p)

(c)
$$\frac{X_1 + X_2 + \ldots + X_N}{\sqrt{N}} \xrightarrow{d} N(0, \sigma^2), \quad \lambda \to \infty.$$
(6 p)

Problem 3

Let X and Y be random variables such that

$$Y \mid X = x \in N(x, a^2) \quad \text{with} \quad X \in N(\mu, b^2).$$

(a) Determine the distribution of the vector $\begin{pmatrix} X \\ Y \end{pmatrix}$. (5 p)

(b) Determine the distribution of $X \mid Y = y$. (5 p)

Problem 4

Let X and Y be two random variables with finite second moment (square-integrable) defined on the probability space $(\Omega, \mathcal{F}, P)$. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$.

(a) Prove (or find a counter-example) that if X and Y are equally distributed (have the same probability distribution) then $E[X \,|\mathcal{G}] = E[Y \,|\mathcal{G}]$ almost surely. (3 p)

(b) Prove that if $E[X \mid Y] = Y$ and $E[Y \mid X] = X$, then $X = Y$ almost surely. (3 p)

(c) Let $X_1, X_2, \ldots, X_n$ be identically distributed and positive random variables. Show that for any $k \leq n$
$$E\left[\frac{X_1 + X_2 + \ldots + X_k}{X_1 + X_2 + \ldots + X_n}\right] = \frac{k}{n}.$$
(4 p)

Problem 5

On the space of random variables on $(\Omega, \mathcal{F}, P)$ we set

$$d(X, Y) := E\left[\frac{|X - Y|}{1 + |X - Y|}\right].$$

(a) Show that d is a metric, that is

$$d(X,X)=0, \quad d(X,Y)=d(Y,X) \quad \text{and} \quad d(X,Y)\leq d(X,Z)+d(Z,Y).$$

(2 p)

Let $X_1, X_2, \ldots,$ be a sequence of random variables on the probability space $(\Omega, \mathcal{F}, P)$.

(b) Show that if $d(X_n, X) \longrightarrow 0$ as $n \to \infty$, then X_n converges to X in probability as $n \to \infty$. (4 p)

(c) Let $X_1, X_2, \ldots,$ be random variables such that

$$P(X_n = 1) = 1 - \frac{1}{n^{1/5}}, \quad P(X_n = n) = \frac{1}{n^{1/5}}, \quad n \geq 2.$$

Determine the limit X such that $d(X_n, X) \longrightarrow 0$ and show that X_n does not converge to X in L^1 i.e. $E[|X_n - X|] \not\longrightarrow 0$ as $n \to \infty$. (4 p)

(*Hint: the function* $f(t) := \frac{t}{1+t}$, $t \geq 0$, *may be useful.*)

KTH Matematik

Avd. Matematisk statistik

Suggested solutions to the exam Wednesday October 26, 2016.

Problem 1

(a) Let $\varphi_{X_1}(t)$ denote the characteristic function of X_1 and $g_N(s) = \exp(\lambda(s-1))$ be the probability generating function of $N \in \text{Po}(a),\ a > 0$. Using the composition formula with characteristic function (5.18) in the Lecture Notes, we have

$$\varphi_{S_N}(t) = g_N(\varphi_{X_1}(t)) = \exp(a(\varphi_{X_1}(t) - 1)).$$

Now, if $S_N \in \text{Po}(b)$, we should have $\varphi_{S_N}(t) = \exp(b(e^{it} - 1))$. Hence,

$$\exp(a(\varphi_{X_1}(t) - 1)) = \exp\left(b(e^{it} - 1)\right),$$

which implies that (note that $0 < b/a < 1$)

$$\varphi_{X_1}(t) = \frac{b}{a}e^{it} + (1 - \frac{b}{a}),$$

that is $X_1, X_2, \ldots$ are all Bernoulli $Be(\frac{b}{a})$-distributed.

(b) Introduce the transformation $(X_1, X_2) := \varphi(R, \Theta),\ R > 0,\ \Theta \in (0, 2\pi)$, defined by

$$\begin{cases} X_1 = R\cos\Theta, \\ X_2 = R\sin\Theta. \end{cases}$$

Then, $R = \sqrt{X_1^2 + X_2^2}$ and $\text{cotan}\Theta = \frac{X_2}{X_1}$. Furthermore, the corresponding Jacobian is

$$J_\varphi := \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\sin\theta \end{vmatrix} = r.$$

We have

$$f_{(R,\Theta)}(r,\theta) = f_{(X_1,X_2)}(r\cos\theta, r\sin\theta)|J_\varphi| = \frac{1}{2\pi} r e^{-\frac{r^2}{2}} = f_\Theta(\theta) f_R(r), \quad r > 0,\ \theta \in (0, 2\pi)$$

where, $f_\Theta(\theta) = \frac{1}{2\pi},\ \theta \in (0, 2\pi)$ which is the density of the uniform distribution over $(0, 2\pi)$ and $f_R(r) = re^{-\frac{r^2}{2}},\ r > 0$ which is the density of the Rayleigh $R(2)$ distribution. Therefore, $\Theta \in \mathcal{U}(0, 2\pi)$ and $R \in R(2)$ and are independent.

Problem 2

The random variable $N \in \text{Po}(\lambda)$ has $E[N] = \text{Var}(N) = \lambda$ and its characteristic function is $\varphi_N(t) = \exp(\lambda(e^{it} - 1))$.

(a) This question can be solved either using Chebychev's inequality or the Cramér-Slutzky's theorem. Using Chebychev's inequality we have, for every fixed $\epsilon > 0$,

$$P\left(\left|\frac{N}{\lambda} - 1\right| > \epsilon\right) = P\left(|N - \lambda| > \lambda\epsilon\right) \leq \frac{1}{\epsilon^2\lambda^2}\mathrm{Var}(N) = \frac{\lambda}{\epsilon^2\lambda^2} = \frac{1}{\epsilon^2\lambda} \to 0,$$

as $\lambda \to \infty$.
Using the Cramér-Slutzky's theorem, then, when $\lambda \to \infty$,

$$\frac{N}{\lambda} \xrightarrow{P} 1 \quad \text{if and only if} \quad \frac{N}{\lambda} \xrightarrow{d} 1 \quad \text{if and only if} \quad \varphi_{\frac{N}{\lambda}}(t) \longrightarrow e^{it}.$$

We have

$$\varphi_{\frac{N}{\lambda}}(t) = \varphi_N\left(\frac{t}{\lambda}\right) = \exp\left(\lambda(e^{i\frac{t}{\lambda}} - 1)\right).$$

But, by Lemma 4.5.1. (in the course lecture notes) we have, for every real x,

$$e^{ix} = \sum_{k=0}^{n} \frac{(ix)^k}{k!} + R_n(x), \tag{1}$$

where

$$|R_n(x)| \leq \min\left(\frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!}\right). \tag{2}$$

In particular, when $x \to 0$, we write

$$e^{ix} = \sum_{k=0}^{n} \frac{(ix)^k}{k!} + o(x^n), \tag{3}$$

where $o(x^n)$ denotes the rest term $R_n(x)$.
Hence, when $\lambda \to \infty$, we obtain (taking n=2)

$$e^{i\frac{t}{\lambda}} = 1 + \frac{it}{\lambda} - \frac{t^2}{2\lambda^2} + o(\frac{t^2}{\lambda^2}). \tag{4}$$

Therefore,

$$\lim_{\lambda\to\infty} \varphi_{\frac{N}{\lambda}}(t) = \lim_{\lambda\to\infty} \exp\left(it - \frac{t^2}{2\lambda} + o(\frac{t^2}{\lambda})\right) = e^{it}.$$

(b) We have to show that $\lim_{\lambda\to\infty} \varphi_{\frac{N-\lambda}{\sqrt{\lambda}}}(t) = e^{-\frac{t^2}{2}}$. Indeed,

$$\varphi_{\frac{N-\lambda}{\sqrt{\lambda}}}(t) = e^{-it\sqrt{\lambda}}\varphi_N\left(\frac{t}{\sqrt{\lambda}}\right) = \exp\left(\lambda(e^{i\frac{t}{\sqrt{\lambda}}} - 1) - it\sqrt{\lambda}\right).$$

We may use (3) to obtain

$$\lambda(e^{i\frac{t}{\sqrt{\lambda}}} - 1) = it\sqrt{\lambda} - \frac{t^2}{2} + \lambda o(\frac{t^2}{\lambda}), \quad \lambda \to \infty.$$

Noting that using (2), $\lambda o(\frac{t^2}{\lambda}) \to 0$ as $\lambda \to \infty$, we finally obtain

$$\lim_{\lambda\to\infty} \varphi_{\frac{N-\lambda}{\sqrt{\lambda}}}(t) = e^{-\frac{t^2}{2}}.$$

(c) Set $S_N = X_1 + X_2 + \ldots + X_N$ and let $\varphi_X(t)$ denote the characteristic function of the i.i.d. variables $X_1, X_2, \ldots$. We have

$$\frac{S_N}{\sqrt{N}} = \frac{S_N}{\sqrt{\lambda}} \Big/ \frac{\sqrt{N}}{\sqrt{\lambda}}.$$

Since $\sqrt{\frac{N}{\lambda}} \xrightarrow{P} 1$ when $\lambda \to \infty$, by the Cramér-Slutzky's theorem it suffices to show that $\frac{S_N}{\sqrt{\lambda}} \xrightarrow{d} N(0, \sigma^2)$ as $\lambda \to \infty$ i.e. we have to show that

$$\lim_{\lambda \to \infty} \varphi_{\frac{S_N}{\sqrt{\lambda}}}(t) = e^{-\frac{t^2\sigma^2}{2}}.$$

Since $N \in \text{Po}(\lambda)$, we have

$$\varphi_{\frac{S_N}{\sqrt{\lambda}}}(t) = \varphi_{S_N}(\frac{t}{\sqrt{\lambda}}) = g_N\left(\varphi_X(\frac{t}{\sqrt{\lambda}})\right) = \exp\left(\lambda\left(\varphi_X(\frac{t}{\sqrt{\lambda}}) - 1\right)\right).$$

Now, since X has finite variance σ^2, we use the fact that, when $\lambda \to \infty$,

$$\varphi_X(\frac{t}{\sqrt{\lambda}}) = 1 + \frac{it}{\sqrt{\lambda}}E[X] - \frac{t^2}{\lambda}E[X^2] + o(\frac{t^2}{\lambda})$$

and noting that $E[X] = 0$ and $E[X^2] = \text{Var}(X) + (E[X])^2 = \sigma^2$, we obtain

$$\lim_{\lambda \to \infty} \lambda\left(\varphi_X(\frac{t}{\sqrt{\lambda}}) - 1\right) = -\frac{t^2}{2}\sigma^2.$$

or

$$\lim_{\lambda \to \infty} \varphi_{\frac{S_N}{\sqrt{\lambda}}}(t) = e^{-\frac{t^2\sigma^2}{2}}.$$

Problem 3

(a) We compute the characteristic function of $\begin{pmatrix} X \\ Y \end{pmatrix}$. We have

$$\varphi_{(X,Y)}(s,t) = E\left[e^{isX+itY}\right] = E\left[E\left[e^{isX+itY}|X\right]\right] = E\left[e^{isX}E\left[e^{itY}|X\right]\right].$$

Since $Y|X \in N(X, a^2)$ we have $E\left[e^{itY}|X\right] = e^{itX - \frac{t^2}{2}a^2}$.
Therefore,

$$\varphi_{(X,Y)}(s,t) = E\left[e^{isX}e^{itX-\frac{t^2}{2}a^2}\right] = e^{-\frac{t^2}{2}a^2}E\left[e^{i(s+t)X}\right].$$

Using the assumption that $X \in N(\mu, b^2)$ we further obtain

$$\varphi_{(X,Y)}(s,t) = \exp\left(is\mu + it\mu - \frac{1}{2}(s^2b^2 + 2stb^2 + (a^2+b^2)t^2)\right).$$

Therefore,

$$\begin{pmatrix} X \\ Y \end{pmatrix} \in N\left(\begin{pmatrix} \mu \\ \mu \end{pmatrix}, \begin{pmatrix} b^2 & b^2 \\ b^2 & a^2+b^2 \end{pmatrix}\right).$$

(b) We have $X|Y = y \in N(\mu_{1|2}, \sigma^2_{1|2})$, where

$$\mu_{1|2} = \mu + \frac{b^2}{a^2+b^2}(y-\mu), \quad \sigma^2_{1|2} = b^2 - \frac{b^4}{a^2+b^2} = b^2\left(1 - \frac{b^2}{a^2+b^2}\right).$$

Note that $\sigma^2_{1|2} > 0$.

Problem 4

(a) Recall the definition of the conditional expectation of an integrable random variable X w.r.t. a sub σ-algebra $\mathcal{G}$. $Z := E[X\,|\mathcal{G}]$ is the unique $\mathcal{G}$-measurable random variable for which

$$E[Z\xi] = E[X\xi], \quad \text{for all bounded } \mathcal{G} - \text{measurable random variables } \xi.$$

But, if X and Y have the same distribution then it is not always true that

$$E[X\xi] = E[Y\xi].$$

In particular, it is not always true that

$$E[E[X\,|\mathcal{G}]\xi] = E[X\xi] = E[Y\xi] = E[E[Y\,|\mathcal{G}]\xi],$$

for all bounded $\mathcal{G}$-measurable random variables ξ. Here is a counter-example. We know that if $X \in U[0,1]$ then $Y := 1 - X \in U[0,1]$. Let ζ be a bounded random variable such that $E[\zeta E[X|\mathcal{G}]] \neq 0$, but $E[\zeta] = 0$ and $\mathcal{G} = \sigma(\zeta)$ i.e. the σ-algebra generated by ζ. If $E[X\zeta] = E[Y\zeta] = E[\zeta] - E[X\zeta]$. Then $0 = E[\zeta] = 2E[X\zeta] = 2E[\zeta E[X|\mathcal{G}]] \neq 0$, leading to a contradiction.

(In a previous version, the suggested solution was wrong.)

(b) To show that $X = Y$ almost surely it suffices to show that $E[(X-Y)^2] = 0$ because both are square-integrable. We have

$$E[(X-Y)^2] = E[X^2 - 2XY + Y^2] = E[X^2] + E[Y^2] - E[XY] - E[XY].$$

But, using the assumption that $X = E[Y|X]$ we obtain

$$E[XY] = E[E[XY|X]] = E[XE[Y|X]] = E[X^2],$$

and using that $Y = E[X|Y]$ we also obtain

$$E[XY] = E[E[XY|Y]] = E[YE[X|Y]] = E[Y^2].$$

Therefore,

$$E[(X-Y)^2] = E[X^2] + E[Y^2] - E[XY] - E[XY] = E[X^2] + E[Y^2] - E[X^2] - E[Y^2] = 0.$$

(c) Set, for $k \leq n$, $S_k = X_1 + X_2 + \ldots + X_k$. Since $X_1, X_2, \ldots, X_n$ are identically distributed it is easy to see that

$$E[X_1|S_n] = E[X_2|S_n] = \ldots = E[X_n|S_n] \quad \text{almost surely.}$$

But, since $S_n = E[S_n|S_n] = \sum_{j=1}^{n} E[X_j|S_n] = nE[X_1|S_n]$, it follows that

$$E[X_j|S_n] = \frac{S_n}{n} \quad \text{almost surely}, \quad j = 1, \ldots, n.$$

This in turns yields that

$$E[S_k|S_n] = \sum_{j=1}^{k} E[X_j|S_n] = \frac{k}{n} S_n \quad \text{almost surely}.$$

Therefore,

$$E\left[\frac{S_k}{S_n}\right] = E\left[E\left[\frac{S_k}{S_n}|S_n\right]\right] = E\left[\frac{E\left[S_k|S_n\right]}{S_n}\right] = \frac{k}{n} E\left[\frac{S_n}{S_n}\right] = \frac{k}{n}.$$

Problem 5

The following properties of the function $t \mapsto f(t) = \frac{t}{1+t}$, $t \geq 0$ turn out useful for solving the problem.

(i) $0 \leq f(t) < 1$ and is strictly increasing since $f'(t) = \frac{1}{(1+t)^2} > 0$.

(ii) f is invertible because it is continuous (in fact differentiable) and strictly increasing.

(iii) $f(t+s) \leq f(t) + f(s)$. indeed, noting that $1+t+s \geq 1+t, 1+s$, we have

$$f(t+s) = \frac{t+s}{1+t+s} = \frac{t}{1+t+s} + \frac{s}{1+t+s} \leq \frac{t}{1+t} + \frac{s}{1+s} = f(t) + f(s).$$

(iii) From (i) and (ii) if follows that if a, b and c are non-negative and satisfy $a \leq b+c$ then $f(a) \leq f(b) + f(c)$.

(a) We have $d(X,X) = E\left[\frac{|X-X|}{1+|X-X|}\right] = 0$. Since $|X-Y| = |Y-X|$ we get $d(X,Y) = d(Y,X)$. Lastly, by the triangle inequality, it holds that $|X-Z| \leq |X-Y| + |Y-Z|$. We may apply (iii) to obtain that

$$\frac{|X-Z|}{1+|X-Z|} \leq \frac{|X-Y|}{1+|X-Y|} + \frac{|Y-Z|}{1+|Y-Z|}.$$

Take expectation to get

$$d(X,Y) \leq d(X,Z) + d(Z,Y).$$

(b) If $d(X_n, X) \to 0$ as $n \to \infty$ we want to show that for every fixed $\epsilon > 0$, $P(|X_n - X| > \epsilon) \to 0$ as $n \to \infty$.
Indeed, using the fact that f is a strictly increasing function and invertible, we have

$$P\left(|X_n - X| > \epsilon\right) = P\left(f(|X_n - X|) > f(\epsilon)\right) \leq \frac{1}{f(\epsilon)} E[f(|X_n - X|)] = \frac{1}{f(\epsilon)} d(X_n, X) \to 0,$$

as $n \to \infty$, noting that since $\epsilon > 0$, $f(\epsilon) > 0$.
(c) We first show that $d(X_n, 1) \to 0$ as $n \to \infty$. Indeed, we have

$$d(X_n, 1) = E\left[\frac{|X_n - 1|}{1+|X_n - 1|}\right] = \frac{n-1}{1+n-1} \frac{1}{n^{1/5}} = \frac{n-1}{n} \frac{1}{n^{1/5}} \to 0,$$

as $n \to \infty$. Hence, $X = 1$.
On the other hand

$$E\left[|X_n - 1|\right] = (n-1)\frac{1}{n^{1/5}} \to +\infty, \quad n \to \infty.$$

Thus, $X_n \not\longrightarrow 1$ in L^1.