UNIVERSITY OF OSLO

Candidate no

Faculty of mathematics and natural sciences

Constituent exam in: STK4540 — Non-life Insurance Mathematics

Day of examination: 12th December 2019

Examination hours: 02:30 pm – 6:30 pm

This problem set consists of 7 pages.

Appendices: None

Permitted aids: Approved calculator

Please make sure that your copy of the problem set is complete before you attempt to answer anything.

This grade chart is only a guideline:

A (85-100) , B (70-85), C (58-70), D (49-57) , E (40-48), F (0-39).

The final grade is set based on an evaluation of the exam as a whole.

Problem 1 (weight 4 points)

(a) The Cramér-Lundberg model is a model for the total claim amount $S(t) = \sum_{i=1}^{N(t)} X_i$ where N is specified to be a homogeneous Poisson process, $\{X_i\}_{i\geq 1}$ are i.i.d. and independent of N. More concretely, we have the following specifications:

- Claims happen at the arrival times $0 \leq T_1 \leq T_2 \leq \ldots$ of a homogeneous Poisson process $N(t) = \#\{i \geq 1 : T_i \leq t\}$, $t \geq 0$.
- The ith claim arriving at time T_i causes the claim size X_i. The sequence $\{X_i\}_{i\geq 1}$ constitutes an i.i.d. sequence of non-negative random variables.
- The sequences $\{T_i\}_{i\geq 1}$ and $\{X_i\}_{i\geq 1}$ are independent. In particular, N and $\{X_i\}_{i\geq 1}$ are independent.

Give 1p if the answer is correct. Give 0.5p if N is chosen to be a homogeneous Poisson process.

(b) The risk process is defined as

$$U(t) = u + ct - S(t), \quad t \geq 0,$$

i.e. initial capital (usually big), plus income from the premiums minus the claims. Define $Z_1 = X_1 - cW_1$, where X_1 is the first claim,

(Continued on page 2.)

$c > 0$ the premium rate and W_1 the inter-arrival time to the first claim. It is enough to look at Z_1 since $\{X_i\}_{i\geq 1}$ and $\{W_i\}_{i\geq 1}$ are i.i.d. A necessary condition to avoid ruin with probability one is

$$E[Z_1] = E[X_1] - cE_[W_1] < 0,$$

meaning that, in average, the claim size "$E[X_1]$" should not exceed the income "$cE[W_1]$".

Give 0.75p if the condition is stated. Give 0.25p for the interpretation.

(c) The Bühlmann model assumes the following:

- The ith policy is described by the pair $(\theta_i, \{X_{i,t}\}_{t\geq 1})$, where the random parameter θ_i is the heterogeneity parameter and $\{X_{i,t}\}_{t\geq 1}$ is the sequence of claim sizes or claim numbers in the policy.
- The sequence of pairs $(\theta_i, \{X_{i,t}\}_{t\geq 1})$, $i = 1, 2, \ldots,$ is iid.
- The sequence $\{\theta_i\}_{i\geq 1}$ is i.i.d.
- Conditionally on θ_i, the sequence $\{X_{i,t}\}_{t\geq 1}$ are independent and their expectation and variance are given functions of θ_i:

$$\mu(\theta_i) = E[X_{i,t}|\theta_i], \quad v(\theta_i) = Var[X_{i,t}|\theta_i].$$

Give 0.5p if the candidate defines the heterogeneity model instead. Give 0.5p if the candidate explains that the point of this model is to have a distribution-free assumption given θ_i. Give 1p if the definition is correct, and 0.25p per assumption.

(d) Observe that

$$e_F(x)\overline{F}(x) = \int_x^\infty \overline{F}(y)dy$$

by the definition of e_F. Then we need to prove that

$$\int_x^\infty \overline{F}(y)dy = e_F(0)e^{-\int_0^x \frac{1}{e_F(y)}dy}.$$

Both sides are positive so applying logarithms we get

$$\log\left(\int_x^\infty \overline{F}(y)dy\right) = \log(e_F(0)) - \int_0^x \frac{1}{e_F(y)}dy.$$

By the fundamental theorem of calculus (on the unbounded interval $(0, \infty)$) we have that both sides are continuously differentiable and by differentiating we have

$$\frac{-\overline{F}(x)}{\int_x^\infty \overline{F}(y)dy} = -\frac{1}{e_F(x)},$$

which is a true identity by the definition of e_F

(Continued on page 3.)

This formula actually shows that, whenever F is continuous and $F(x) > 0$ for $x > 0$ then e_F and F uniquely determine each other. The mean excess function is used to determine whether a random variable has heavy or light tails being the criterion: $\lim_{u\to\infty} e_F(u) = 0$ light tails and $\lim_{u\to\infty} e_F(u) = \infty$ heavy tails.

Give 0.5p for the proof and 0.5p for the explanation related to small/large claims. Give also points if they assume F is differentiable to prove the identity.

Problem 2 (weight 4 points)

This exercise is worth (a) 0.5p, (b) 0.5p, (c) 1p, (d) 0.5p, (e) 0.5p and (f) 1p.

(a) we can compute $E[X_1 \mathbf{1}_{\{X_1 \le L\}}]$ by conditioning on X_1:

$$E[X_1 \mathbf{1}_{\{X_1 \le L\}}] = E[X_1 | X_1 \le L] P(X_1 \le L) = \frac{L+1}{2}\frac{L}{M}.$$

Give 0.5p if correct.

(b) Again, by conditioning

$$\begin{aligned} E[e^{i\theta X_1 \mathbf{1}_{\{X_1 \le L\}}}] &= E[e^{i\theta X_1} | X_1 \le L] P(X_1 \le L) + 1 P(X_1 > L) \\ &= \varphi_L(\theta)\frac{L}{M} + \frac{M-L}{M}. \end{aligned}$$

Give 0.5p if correct.

(c) Since characteristic functions characterize distributions we will compute both and check whether they are equal. In general, for a compound Poisson process of the form $H(t) = \sum_{i=1}^{M(t)} Y_i$ being M Poisson with intensity μ and Y_i i.i.d., one has

$$E[e^{i\theta H(t)}] = E\left[e^{i\theta \sum_{i=1}^{M(t)} Y_i}\right] = E[\varphi_{Y_1}(\theta)^{M(t)}] = e^{-\mu t \left(1-\varphi_{Y_1}(\theta)\right)},$$

where φ_{Y_1} denotes the characteristic function of Y_1.

So, on the one hand

$$E[e^{i\theta S(t)}] = e^{-\lambda t \left(1-\varphi_{X_1 \mathbf{1}_{\{X_1 \le L\}}}(\theta)\right)} = e^{-\lambda t \frac{L}{M}(1-\varphi_L(\theta))},$$

where the latter step follows from (c). On the other hand,

$$E[e^{i\theta \widehat{S}(t)}] = e^{-\widehat{\lambda} t (1-\varphi_L(\theta))} = e^{-\lambda t \frac{L}{M}(1-\varphi_L(\theta))}.$$

Since both characteristic functions coincide, the random variables $S(t)$ and $\widehat{S}(t)$ have equal distributions for every $t \ge 0$. Obs: this does

(Continued on page 4.)

not mean, that $S \sim \widehat{S}$ as a random variable in an infinite dimensional space, this is the next item.

Give 0.5p per computation and 1p if everything is correct independently of the method. Subtract 0.25p for major mistakes.

(d) It is enough to prove it for $n = 2$. We have

$$\begin{aligned} E\left[e^{i(\theta_1 S(t_1)+\theta_2 S(t_2))}\right] &= E\left[e^{i(\theta_1+\theta_2)S(t_1)+\theta_2(S(t_2)-S(t_1))}\right] \\ &= E\left[e^{i(\theta_1+\theta_2)S(t_1)}\right] E\left[e^{i\theta_2(S(t_2)-S(t_1))}\right] \\ &= E\left[e^{i(\theta_1+\theta_2)S(t_1)}\right] E\left[e^{i\theta_2(S(t_2-t_1))}\right] \\ &= E\left[e^{i(\theta_1+\theta_2)\widehat{S}(t_1)}\right] E\left[e^{i\theta_2(\widehat{S}(t_2-t_1))}\right] \\ &= \dots \\ &= E\left[e^{i(\theta_1\widehat{S}(t_1)+\theta_2\widehat{S}(t_2))}\right]. \end{aligned}$$

Give 0.5p if it is correct.

(e) We have

$$\inf_{t\geq 0} U(t) = \inf_{t\geq 0} (u + ct - S(t)).$$

Bankrupcy can only happen at the arrival times T_n of the claims since the premium income is a linear function. Hence,

$$\inf_{t\geq 0} U(t) = \inf_{n\geq 0} (u + cT_n - S(T_n)).$$

But now we have a countable set of $S(T_n)$ so we can say $S(T_n) \sim \widehat{S}(T_n)$ and the claim follows. Give 0.5p if it is correct.

(f) Let $Z_1 = \widehat{X}_1 - c\widehat{W}_1$ being $W_1 \sim exp(\widehat{\lambda})$. Then the Lundberg coefficient is the positive root of the equation $m_{Z_1}(h) = 1$. But $m_{Z_1}(h) = m_{\widehat{X}_1}(h) m_{\widehat{W}_1}(-ch)$. Using the definitions of moment generating functions for both variables we obtain

$$\frac{e^h - e^{(L+1)h}}{L(1-e^h)} \frac{\widehat{\lambda}}{\widehat{\lambda} + ch} = 1.$$

Now using $e^h \approx 1 + h$ we have

$$\frac{1 + h - 1 - (L+1)h}{-Lh} \frac{\widehat{\lambda}}{\widehat{\lambda} + ch} \approx 1.$$

Further,

$$\frac{\widehat{\lambda}}{\widehat{\lambda} + ch} \approx h \iff ch^2 + \widehat{\lambda} h - \widehat{\lambda} = 0 \iff r \approx \frac{-\widehat{\lambda} + \sqrt{\widehat{\lambda}^2 + 4c\widehat{\lambda}}}{2c} > 0.$$

(Continued on page 5.)

Hence, the probability of ruin is bounded by

$$\psi(u) \leq e^{-ru}$$

Give 0.5p for proving the equation and 0.5p for finding the approximate root.

Problem 3 (weight 4 points)

(a) The posterior density satisfies the following proportionality

$$f_p(y|\vec{X} = \vec{x}) \propto \left(\prod_{i=1}^{n} P(X_1 = x_i | p = y) \right) f_p(y).$$

Then we can write the content of each expression that only depends on y,

$$\begin{aligned} f_p(y|\vec{X} = \vec{x}) &\propto \left(\prod_{i=1}^{n} y(1-y)^{x_i - 1} \right) y^{\alpha - 1}(1-y)^{\beta - 1} \\ &= y^{n+\alpha-1}(1-y)^{\sum_{i=1}^{n} x_i - n + \beta}. \end{aligned}$$

The latter is a function proportional to the density function of a beta distributed random variable with parameters $\overline{\alpha} = \alpha + n$ and $\overline{\beta} = \beta + \sum_{i=1}^{n} x_i - n$. Observe that $x_i \geq 1$ for all i so $\sum_{i=1}^{n} x_i - n \geq 0$. Give 1p for writing down the formula for the conditional denisty. Give 2p if it is correct.

(b) The Bayes estimator is given by the posterior mean, i.e.

$$\begin{aligned} \widehat{\mu}_B(p) &= E[\mu(p)|\vec{X}] \\ &= E\left[\frac{1}{p}|\vec{X}\right] \\ &= \int_0^1 \frac{1}{y} f_p(y|\vec{X} = \vec{x})dy \\ &= \frac{\Gamma(\overline{\alpha} + \overline{\beta})}{\Gamma(\overline{\alpha})\Gamma(\overline{\beta})} \int_0^1 \frac{1}{y} y^{\overline{\alpha}-1}(1-y)^{\overline{\beta}-1}dy \\ &= \frac{\Gamma(\overline{\alpha} + \overline{\beta})}{\Gamma(\overline{\alpha})\cancel{\Gamma(\overline{\beta})}} \frac{\Gamma(\overline{\alpha} - 1)\cancel{\Gamma(\overline{\beta})}}{\Gamma(\overline{\alpha} - 1 + \overline{\beta})} \underbrace{\frac{\Gamma(\overline{\alpha} - 1 + \overline{\beta})}{\Gamma(\overline{\alpha} - 1)\Gamma(\overline{\beta})} \int_0^1 y^{\overline{\alpha}-2}(1-y)^{\overline{\beta}-1}dy}_{=1} \\ &= \frac{\Gamma(\overline{\alpha} + \overline{\beta})}{\Gamma(\overline{\alpha} + \overline{\beta} - 1)} \frac{\Gamma(\overline{\alpha} - 1)}{\Gamma(\overline{\alpha})}. \end{aligned}$$

Finally using the fact that $\frac{\Gamma(z)}{\Gamma(z-1)} = z - 1$ and $\frac{\Gamma(z'-1)}{\Gamma(z')} = \frac{1}{z'}$ we conclude that

$$\widehat{\mu}_B(p) = \frac{\alpha + \beta + \sum_{i=1}^{n} X_i - 1}{n + \alpha}.$$

(Continued on page 6.)

If we write

$$\widehat{\mu}_B(p) = \frac{\alpha + \beta - 1}{n + \alpha} + \frac{\overline{X}_n}{1 + \alpha/n}$$

we can apply the (conditionally on p) strong law of large number to conclude that

$$\widehat{\mu}_B(p) \to E[X_1|p] = \mu(p)$$

almost surely as $n \to \infty$ and hence $\widehat{\mu}_B$ is strongly consistent.

Give 0.5p for saying that the Bayes estimator is the posterior expectation. Give 1p for the Bayes estimator and 1p for the strong consistency. Give 2p if everything is correct and subtract points if there are major mistakes.

Problem 4 (weight 2 points)

Using the formula for the CLM estimate given in the exercise we can fill in the run-off triangle and obtain

Cumulative claims loss settlements		Development year			
		0	1	2	3
Claims occurrence year	2016	967	1143	1546	1634
	2017	1054	1365	1587	1677
	2018	1287	1477	1845	1950
	2019	1734	2089	2609	2758
CLM estimator for claims loss settlement factor			1,2047	1,2492	1,0569

Figure 1: Observed and estimated cumulative payments.

The technical provisions for year, say 2020, are the sum of the incremental claims loss settlements. This corresponds to the incremental settlements placed in the diagonal corresponding to year 2020, that is $(1677 - 1587) + (1845 - 1477) + (2089 - 1734) = 813$. Repeating this procedure for each year (increments in each diagonal) we obtain the following table:

Calendar year	Estimated claims loss settlement amounts
2020	813
2021	626
2022	149

Figure 2: Technical provisions for future years

(Continued on page 7.)

which are the provisions for the years 2020, 2021 and 2022.

Give 1p if the candidate has computed the estimated claims loss settlements. Give 1p for the estimated loss settlement amounts for years 2017 and 2019.

The final point sum is a number X between 0 and 14. The grade is then computed as a number between 0 and 100 as follows

$$\text{Grade} = \frac{100}{14}X.$$