ACTUARIAL RESEARCH CLEARING HOUSE
1993 VOL. 2

TEACHING RISK THEORY

Wojciech Szatzschneider
School of Actuarial Sciences
Anahuac University and
Department of Mathematics
Universidad Autónoma Metropolitana Iztapalapa
Mexico City

I lecture an undergraduate course on Risk Theory at Anahuac University, School of Actuarial Sciences, and from time to time, a graduate one in the Mathematical Department of Universidad Autónoma Metropolitana.

In the present comment I would like to express some opinions stemming mostly from my experience at Anáhuac University. One year ago Mexican students presented, for the first time, the 151 examination. They obtained good results.

For the most part, this comment is siding in the direction of classical risk theory, this being obligatory according to Society of Actuaries syllabus. In **Actuarial Mathematics** a great job has been done resulting in a textbook accessible to average students. Consequently, the main goal of my comments is to make the next edition of **Actuarial Mathematics** slightly better in the part involving Risk Theory.

My first comment, which crops up from our experience, is the following: the $o(f(t))$ formalism and notation should be included. For example $f(t)= o(t)$, $t \to 0$, if $\frac{f(t)}{t} \to 0$

This formalism is similarly proving to be very helpful in the 100 and 140 examinations. Many problems in the 100 examination may be solved without the laborious and mechanical L'Hospital's rule. The $o(t)$ formalism is even more helpful, in the 140 examination. Unfortunately, the textbook by Kellison does not contain this kind of approximation. It is the general opinion of our academic staff—especially that of Luisa Ares, the professor who is teaching the theory of interest, Miguel Angel Flores, calculus teacher, and mine, that this kind of infinitesimal calculus should be included for actuarial students. That is what we are doing here, making easier estimations of error.

For example, the proof of the central limit theorem is easier to explain with the little-oh formalism. We will see an application of little-oh, when talking about Theorem 11.4. (from **Actuarial Mathematics**).

Another general comment of mine is that the texbook is focused too much in the use of moment generating function. Proofs are not always easier using the m.g.f. technique, and moreover, the general idea often escapes, making the m.g.f. technique somewhat mindless.

Now I would like to discuss some details of the presentation of Risk Theory in A.M.

Chapter 1.

1.- The text concerning

$$G = \frac{\log M_X(\alpha)}{\alpha}$$

and

$$H = \frac{\log M_X(\alpha_I)}{\alpha_I}$$

should continue just like in Gerber's book: *An Introduction to Mathematical Risk Theory*.
Many students are asking: if $\alpha > \alpha_I$ is the policy feasible?

With the trivial modification, one does not need the concept of the Essher transformation.

$\frac{M''(t)}{M(t)}$ is, for any given t, the second moment of some r.v. with density function

$$\frac{e^{tx} f(x)}{\int e^{tx} f(x) dx}$$

2.- At least one example should be attached to the text showing that, for the utility function $\log\omega$, the premium G is decreasing with ω, and an explanation why the premium increases with the wealth in the case of quadratic utility. It seems that this last property is a very particular one of the quadratic utility function. As a matter of fact, it is due to increasing curvature of this function.
Thus the "unexpected property" results more from geometrical than from algebraical properties.

Chapter 2.

1.- Page 30 is badly edited. The concepts are easy, but the presentation is dull and readers lose patience.

2.- In my class I also solve the example from 2.2 in a third way, conditioning with respect to three values of I.

$$I = \begin{cases} 0 \text{ if the claim does not occur} \\ 1 \text{ if it does and its amounts} < 2{,}000 \\ 2 \text{ if it occurs and is} = 2{,}000 \end{cases}$$

It may be not so natural to choose more than 2 values for I, but it is a good exercise to solve the problem this way, practicing conditioning techniques.

3.- Talking about the central limit theorem: (page 36) "... The central limit theorem does extend to sequences of nonidentically distributed random variables", some additional comment should be added like. "If anyone of r.v. X_i is not too big comparing with the others". There are references, but the sentence sounds too categorical.

Chapter 11.

1.- The introduction to this chapter makes one think that only because of single period the interest rate is not included; but in the following chapter about the multi-periods, the interest rate is not included either, so the argument does not stand.

2.- One task of a teacher in a classroom should be to link the risk theory with remaining courses. A good example could be 11.4, which may be solved by "tying ends" in bayesian statistics. Prior $\Gamma_{(\alpha,\beta)}$ distribution and Poisson random sample will produce posterior $\Gamma_{(\alpha+x,\beta+1)}$ distribution. From this observation it is very easy to solve the example.

3.- Th.11.1. It would be nice to come back to this theorem in chapter 12 or 13 with more explicit and natural proof based on the fact that if $Z_1, Z_2, \ldots, Z_n$ are waiting times, Z_i for a claim from the policy i, and Z_i are independent exponential r.v. then, if a claim comes, it is from the policy i with probability

$$\frac{q_i}{q_1 + \ldots + q_n} = \frac{q_i}{q}$$

The argument which could lead to this relation can not be entirely rigorous. One should use the strong Markov property, but we follow a slightly informal way of proofs following A.M.
Now for Poisson Process; if a claim X occurs,

$$X = \begin{cases} X_1 \text{ with prob. } \frac{q_1}{q} \\ \ldots \\ X_n \text{ with prob. } \frac{q_n}{q} \end{cases}$$

So the theorem results for Poisson process for each t. Therefore it results for Poisson distribution.

On the other hand, there is a close relation between Theorem 11.1 and 13.3. We will talk about this later.

4.- In the proof of Theorem 11.2, a new guest appears which should not be invited, namely, the multidimensional moment generating function which usually lies beyond the elementary probability courses (it clearly lies beyond the 110 examination). We simply do not need it.

The proof that $\mathcal{N}_i$ follows Poisson distribution law is quite clear. (Using, e.g., standard m.g.f. or the law of total probability). So it remains to prove the independence of $\mathcal{N}_i$. But this is a consequence of the following easy argument.

Let $\mathcal{N} = \sum \mathcal{N}_i$, and $n = \sum n_i$. Then

$$\begin{aligned} & P(\mathcal{N}_1 = n_1, \mathcal{N}_2 = \mathcal{N}_2, \ldots, \mathcal{N}_m = n_m) \\ = \ & P(\mathcal{N}_1 = n_1, \mathcal{N}_2 = n_2, \ldots, \mathcal{N}_m = n_m, \mathcal{N} = n) \\ = \ & P(\mathcal{N}_1 = n_1, \mathcal{N}_2 = n_2, \ldots, \mathcal{N}_m = n_m \mid \mathcal{N} = n)\, P(\mathcal{N}=n). \end{aligned}$$

The conditional distribution in multinomial (more obvious from our interpretation of Theorem 11.1) and $\mathcal{N}$ has Poisson distribution with parameter λ. So we have

$$P(\mathcal{N}_1 = n_1, \mathcal{N}_2 = n_2, \ldots, \mathcal{N}_m = n_m) = \frac{n!}{n_1! n_2! \ldots n_m!} \pi_1^{n_1} \pi_2^{n_2} \ldots \pi_m^{n_m} \frac{\lambda^n}{n!} e^{-\lambda}$$

$$= \frac{(\lambda\pi_1)^{n_1}}{n_1!} e^{-\lambda\pi_1} \frac{(\lambda\pi_2)^{n_2}}{n_2!} e^{-\lambda\pi_2} \ldots \frac{(\lambda\pi_m)^{n_m}}{n_m!} e^{-\lambda\pi_m}$$

$$= P(\mathcal{N}_1 = n_1)\ P(\mathcal{N}_2 = n_2) \ldots P(\mathcal{N}_m = n_m)$$

5.- Th. 11.3. I consider more important and more didactic to show that the proof relies on the connection of this theorem to the classical central limit theorem instead of proposed technical proof. For the sake of applications it is enough to consider λ and r integers in a, and b, respectively. In this case almost nothing remains to be done. Anyway, it is quite classical that if $P(Z_n \leq x) \to \Phi(x)$, and if $\alpha_n \to 1$, $P(|Y_n| > \epsilon) \to 0$ (Y_n tends to zero in probability), then $P(\alpha_n Z_n + Y_n \leq x) \to \Phi(x)$. This fact is called by statisticians "Slutsky's theorem". This theorem lies beyond the 110 examination. But even some kind of intuitive reasoning would be sufficient to understand the problem.

Now the theorem 11.3a results immediately. Setting $[\cdot]$ for the integer part,

$$\frac{S_\lambda - \lambda p_1}{\sqrt{\lambda p_2}} \ \underline{\underline{D}} \ \frac{S_{[\lambda]} + S_{\lambda-[\lambda]} - [\lambda]p_1 - (\lambda - [\lambda])\, p_1}{\sqrt{[\lambda]p_2}\sqrt{\frac{\lambda}{[\lambda]}}}$$

($S_{[\lambda]}$ and $S_{\lambda-[\lambda]}$ being independent).

The proof of 11.3b. results with trivial changes.

(Note that both compound Poisson and compound negative binomial disitributions is "closed under convolution", considering here common P(x) and in the second case alse common p, and

$\frac{S_{\lambda - [\lambda]}}{\sqrt{[\lambda]}} \to 0$ in probability)

6.- With respect to the translated gamma distribution (p.336), it seems easier to consider $X \sim \Gamma(\alpha,\beta)$ and $S = x_0 + X$ the translated gamma r.v.

The remark on the page 337 is not another limit theorem but results immediately from the central limit theorem and our proof of the Theorem 11.3.

7.- Theorem 11.4 This is a classical example of how the use of little-oh calculus can simplify the proof. In the Taylor's formula (from the textbook), too many terms are involved. It is enough to use $f(x) = f(0) + f'(x)\,x + o(x)$ which, as a matter of fact, results from the existence of the first derivative.

Now $\frac{q(k)}{p(k)} = k\,\alpha$, where $\alpha = \frac{q(1)}{p(1)}$. Since $E(S_k) = rk\,p_1\alpha$

$M_X\left(t/E(S_k)\right) = 1 + \frac{p_1\,t}{r\,k\,p_1\,\alpha} + o\left(\frac{1}{k}\right)$. We took for granted the existance of $M_X(t)$

within an interval $(-\infty,\gamma)$, $\gamma > 0$, and this clearly implies the existance of $M'_X(0)$, and $E(|X|^k)$ for any k. The same assumption is made in A.M., but it does not appear until the chapter 13.

Now, m.g.f. of $S_k/E\left(S_k\right)$ is $\left[\frac{p(k)}{1-q(k)\,M_X\left(t/E(S_k)\right)}\right]^r$

We will calculate

$$\frac{1-q(k)\,M_X\left(t\,/\,E\,(S_k)\right)}{p(k)} = \frac{1-\left(1-p(k)\right)}{p(k)} - \frac{q(k)\,t}{p(k)\,r\,k\,\alpha} + k o\left(\frac{1}{k}\right) \to \left(1-\frac{t}{r}\right)$$

Therefore the limit of the m.g.f. of $\frac{S_k}{E(S_k)}$ is $\left(\frac{r}{r-t}\right)^r$ as k tends to infinity. This proof allows an immediate generalization:

Instead of asking for

$\frac{q(k)}{p(k)} = k\,\alpha$ we might assume that $\lim_{k\to\infty} \frac{q(k)}{p(k)}\frac{1}{k} = \alpha$

It is quite important that students realize that unlike in the central limit theorem, here only the first moment is involved.

8.- I propose the following change in the formulation of problem 11.23. If X_i are individual claims in the second model, then individual claims Z_i in the first model are $Z_i = X_1 + X_2 + X_{\mathcal{N}}$, X_i, $\mathcal{N}$ being independent random variables and $P(\mathcal{N}=k) = \frac{q^k}{-k \ln p}$. In this presentation the structure of the first model is clearer.

Chaper 12

1.- The global method a) in the page 348 should read: "... We specify the distribution of $\mathcal{N}(t+h) - \mathcal{N}(t)$ given the past history $\mathcal{N}(s)$, $s \le t$. This distribution may depend, or not depend, of the values of $\mathcal{N}(s)$, $s \le t$"

2.- The method b) is not easy to understand. More explanation is needed in order to "catch" the loss of generality by this method.

3.- 12.3 gives a rather strange example $x^2 e^{-x}$. I consider it much more important that the method of adjustment coeficient fails in very important examples like Pareto or Lognormal. At least some references should be given in the case of "heavy tail" claims.

4.- I propose the following, slightly modified version of the proof of Theorem 12.1. Similar to the A.M. proof, we avoid the formal treatment with martingales.

Start with the lemma (also useful in the proofs of Theorems 12.2 and 12.3);
If $X_n \ge 0$, $E(X_n) = f(n) \to \infty$, and $\mathrm{Var}\left(\frac{X_n}{f(n)}\right) \to 0$, then $E\left(e^{-X_n}\right) \to 0$

Proof of the lemma

Let $0 < \alpha < 1$ be any number. Now we have

$$0 \le E\left(e^{-X_n}\right) = E\left(e^{-X_n} \mid X_n \ge \alpha f(n)\right) P(X_n \ge \alpha f(n)) + E\left(e^{-X_n} \mid X_n < \alpha f(n)\right) P\left(X_n < \alpha f(n)\right)$$

$$\le e^{-\alpha f(n)} + P\left(\mid X_n - f(n) \mid > (1-\alpha) f(n)\right) < e^{-\alpha f(n)} + \frac{\mathrm{Var}\,(X_n)}{(1-\alpha)^2 \left(f(n)\right)^2} \to 0.$$

Proof of the theorem.

$$E\, e^{-RU(t)} = e^{-Ru}\, e^{-Rct}\, E\left(e^{RS(t)}\right) = e^{-Ru}\, e^{-Rct}\, e^{\lambda t\,[M_X(R)-1]}$$

$= e^{-Ru}\, e^{-Rct + \lambda t\,[M_X(R)-1]} = e^{-Ru}$, because R is the adjustment coefficient. On the other hand,

$$E\, e^{-RU(t)} = E\,(e^{-RU(t)} \mid T \le t)\, P\,(T \le t) +$$
$$E\,(e^{-RU(t)} \mid T=\infty)\, P\,(T=\infty) + E\,(e^{-RU(t)} \mid t < T < \infty)\, P\,(t < T < \infty)$$

Now, as $t \to \infty$, $P\,(t < T < \infty) \to 0$, and $E\,(e^{-RU(t)} \mid t < T < \infty) < 1$. So the third term vanishes as t tends to infinity.

If $T = \infty$, $R\,U(t)$ is non-negative

$E\left(R\,U(t)\right) = R\,(u + \alpha\, t)$; $\mathrm{Var}\left(R\,U(t)\right) = R^2\,\beta^2\, t$,

where $\alpha = c - \lambda\, p_1$, $\beta^2 = \lambda\, p_2$.

So we can apply the lemma to the second term, which vanishes as $t \to \infty$. Therefore, it suffices to show that

$E\left(e^{-RU(T)} \mid T \le t\right) = E\left(e^{-RU(t)} \mid T \le t\right)$ and let t tend to infinity.

We will prove first the equality

$$E\left(e^{-R\,U(t)} \mid T = \tau \le t\right) = E\left(e^{-RU(\tau)} \mid T = \tau \le t\right)\, e^{-Rc\,(t-\tau)}.\ E\, e^{RS(t-\tau)}$$

Note that $U(t) = U(\tau) + c(t - \tau) - [S\,(t) - S\,(\tau)]$. We can separate the expectations because $U(\tau)$, given $T=\tau \le t$, and $S(t) - S(\tau)$ are independent. Moreover $S(t)$ has stationary and independent increments, so the equality results easily.

Now, as before, R being the adjustment coeficient implies that

$$E\left(e^{-RU(t)} \mid T=\tau\leq t\right) = E\left(e^{-RU(\tau)} \mid T=\tau\leq t\right). \text{ Therefore}$$

$$E\left(e^{-RU(t)} \mid T\leq t\right) = E\left(e^{-RU(T)} \mid T\leq t\right)$$

5.- The example 12.4 results immediately due to "lack of memory" for exponential distribution (which as a matter of fact, is proved in the textbook). Nevertheless, this lack of memory should be mentioned in an explicit form.

6.- The proof of Theorem 12.4. is too technical and one does not get the point.
What good does it to do the important concept of penalty if it is not used in the sequel and students have found the proof unreadable?

I consider clearer and easier to explain Theorems 12.4 and 12.5 (which are closely related) using the renewal theory, as in the classical texbooks by Feller, Karlin and Taylor, Bűlhmann and Grandell.

Let $\Phi(z) = 1 - \psi(z)$. Now $\Phi(z)$ satisfies:

$$\Phi(z) = \int_0^\infty \lambda\, e^{-\lambda\tau}\left(\int_0^{z+c\tau} \Phi(z+c\tau-x)\, dF(x)\right) d\tau\ , \ z>0$$

Later on, following e.g. Grandell p. 5,

$$\Phi(u) = \Phi(o) + \frac{\lambda}{c}\int_0^u \Phi\,(u-z)\left(1-F(z)\right) dz.$$

By the way, I do not understand the unanimity of these classical texts of claiming that $\Phi(u)$ is differentiable.

The equation

$\Phi'(u) = \frac{\lambda}{c}\,\Phi(u) - \frac{\lambda}{c}\int_0^u \Phi(u-z)\, d\,F(z)$ <u>cannot</u> <u>be</u> <u>satisfied</u> <u>everywhere</u> if $F(z)$ suffers jumps.

There is a counterexample, in A.M. (example 12.22.) In points of jumps of the distribution function F, Φ' would suffer jumps, but it is impossible. (the derivative does not need to be continuous, but all discontinuities are essential). So the solution of the problem 12.22 (p. 613) should read a) ... $0 < u < 1$

The robust result $\psi(0) = \frac{1}{1+\theta}$ might be obtained in the same way as in Grandell's text or, in general, the Theorem 12.4 results straightforward using general theory of renewal processes. The total loss M is smaller than x, if and only if there is no ruin with initial surplus x. M, the maximal loss with infinite horizon, might be seen in the following way: every time when surplus is negative we point the respective quantity out, and start with zero surplus.

Now the accumulated loss satisfies $P(M \leq x) = \Phi(x)$, so the improper density $\frac{\alpha}{c}\left(1-F(y)\right)$ represents what we were seeking. (We may assume initial surplus equal to be zero)

Theorem 12.5 follows immediately from (12.5.4).

Now, to solve the problem 12.22 we simply "solve"

$$\Phi(t) = 1 - \frac{\lambda}{c} + \frac{\lambda}{c} \int_0^{\min(1,t)} \Phi(t-x)\,dx$$

Chapter 13.

1.- The formula $\frac{q_i}{q_1 + \dots + q_n}$ needs more explanation

First of all, why Bayes? It rather results from the definition of conditional probability. Now, if q_i are small, then the probability of two or more claims is smaller only if there are not too many q_i's. Nevertheless, we could treat the problem in a given time interval [0,T], and if $\sum_{i=1}^{n} q_i > 1$, or $\sum q_i q_j$ is of the same or geater order, we could split the original interval in 2 or many equal segments and look for n such that $\sum \frac{q_i}{n} \gg \sum \frac{q_i}{n}\frac{q_j}{n}$. ($\gg$ means here much bigger)

We have assumed that the moment of arrival of each claim given that it occurs is U[0,T] distributed. Now the best way to get the result seems to be the following: P (claim i | claim in t) =

$$\lim_{h \to 0} P\left(\text{claim } i \mid \text{claim in } [t, t+h)\right) = \lim_{h \to 0} \frac{q_i \frac{h}{T}}{q_1 \frac{h}{T} + \dots + q_n \frac{h}{T} + o(h)}$$

Almost the same proof proceeds for the exponential waiting time for claims. (Compare with the comment about th.11.1).

2.- In example 13.11 (Proportional reinsurance with θ (should read ξ) = 100 per cent, which is the same as θ, should be rather called "friendly cooperation" instead of reinsurance. The chapter 13 is called "applications", excess - of - loss coverage with $\xi = \theta$ would be rather difficult to get. At least, it should be discussed.

3.- In page 391 - the sentence "This implies that there is not a positive root to (12.3.1) and that the ruin is certain" is confusing. It is valid in the context, but less careful students might take it to be "the general truth".

Acknowledgments.

I would like to express my gratitud to the referee.
Thanks to him I made several improvements to the text.