CASUALTY RATE PREDICTION FOR OIL TANKERS

Douglas McKenzie

Casualty Rate Prediction for Oil Tankers

Douglas McKenzie

A model of oil tanker casualties is presented which permits an expected casualty rate for each tanker to be calculated based on its age and casualty history. These expected rates are shown to be good predictors of both the actual casualty experience and the probability of total loss. The model is based on the findings that 1) the casualty behavior of an individual tanker follows a Poisson distribution and 2) the Poisson parameters for all tankers of a specific age follow an exponential distribution. As a result, Bayes' theorem can be used to calculate each tanker's expected casualty rate given its age and casualty history.

Ocean Marine Insurance

A brief summary of ocean marine premium-setting practices is given to provide context for the risk model presented.[1]

A five-year average of claims is used as an estimate of the financial risk associated with *small partial losses* (eg. less than $100,000) of a particular owner's fleet of ships. A fleet of five ships with five years of claims history is often considered to be self-rating for this component of the ***hull & machinery insurance***.

Premiums also have to be established for *large partial losses* and *total losses*. These events are much less frequent, so, for most owners, having just a few ships, the rates are established using industry-wide statistics. These general rates are then multiplied by factors which reflect higher or lower relative risk of a particular fleet as compared to the larger fleet from which the statistics were taken. These factors, called 'relativities', are developed for age, size, trade (ie. routes traveled), flag (ie. nationality of registered owner) and anything else that the underwriter believes might affect the risk of loss.

Hull & machinery insurance is generally carried by commercial insurers so the attempt to define a specific fleet's (ie. owner's) level of risk is expected. The underwriting cycles in ocean marine insurance, however, are pronounced hence the actual premiums that are charged may not always directly reflect that risk.[2] After several years of disappointing underwriting results, though, the early '90s have seen rates, deductibles and exclusions all increase dramatically. In addition to the overall rise in premiums, underwriters are making unprecedented efforts to identify 'substandard' vessels that require special attention even to the point of contracting ship inspections.

The liability side of ocean marine insurance, called ***P&I insurance*** for protection and indemnity, is largely handled by mutualized shipowner groups known as the 'P&I clubs'. 'Advance calls' are prepaid by the shipowners early in the year and 'supplementary calls' are made if the aggregate of advance calls do not cover all the claims that year. Unlimited coverage is provided except for oil pollution with a $500 million limit with another layer of $200 million of protection available commercially.

It is, perhaps, somewhat less clear than with hull insurers how the P&I clubs allocate the total calls required to the specific fleets at risk since the statistics available are even more limited. Access to supplementary calls may make the question less compelling than with hull insurance in the commercial arena. Underwriters for the commercial layer of oil pollution coverage, though, *are* trying to improve selectivity. Just this year, for example, London underwriters agreed on a schedule of rates depending on age and hull design features as indicators of risk.

It is also pertinent to note that at this time, and for the foreseeable future, freight rates are generally depressed and cannot support the aging fleet's replacement needs. Many in the industry feel that insurers could help this situation by pricing insurance for the substandard ships high enough to drive them into the scrap yards.

It appears, from the description above, that improved estimates of risk could be of use to the industry at this time. This paper presents a new model of tanker risk which combines the two types of risk estimates currently being used into a single, consistent framework based on 'reported casualties'. The two types of estimates that are combined are:

> Five-year averaging of claims within a fleet for the more frequent, small losses and
>
> Statistically derived rates for the rare, large losses calculated by looking across all the fleets.

The model presented represents a first attempt at this consolidation and much work remains to be done.

Contents of the Paper

- 'Reported casualties' are introduced as a surrogate for actual claims.
- The statistics of these reported casualties are then described.
- The method used to calculate an expected casualty rate for each ship, using the statistics, is described.
- Modifications to the basic model are briefly discussed.
- Comparison of the calculated casualty rates with actual casualty experience is made for 1991 and 1992.
- The expected casualty rates are used to predict the probability of total loss.
- Areas in which further research is needed are discussed.

Reported Casualties

Combining the two types of risk estimates requires a new variable to overcome the general unavailability of claims information. Even marine underwriters may not have reliable five-year claims information if the fleet being considered is new to them.

Lloyd's List, a daily newspaper published by Lloyd's of London Press, provides a suitable variable. The List reports casualties incurred by all types of ships all around the world. These casualties, reported by the network of Lloyd's agents following ship activities all around the world, are used as a surrogate for claims.

The New York-based Tanker Advisory Center has compiled a unique database which includes every oil tanker casualty reported in Lloyd's List since 1964. The Center has kindly made this database available to Pyramid Systems to make the analysis reported here possible.

This analysis focuses on 2,500 privately-owned oil tankers which incur between 350 and 450 casualties each year that are reported in the List. These casualties range in severity from plugged fuel lines to total loss. They do not usually have financial impact associated with them. The overall frequency of these casualties (1 per ship every 5-7 years) is seen to lie between the more frequent small insurance claims and the less frequent large claims.

There are a few points to make about these casualties before describing their statistical behavior:

- Virtually all 'serious' incidents undoubtedly appear as reported casualties. This common sense expectation is supported by the experience of government-owned vessels. These ships have substantially fewer reported casualties than privately-owned vessels, however, they have essentially the same number of serious casualties. This is probably due to mechanisms which can render the small casualties incurred by government ships invisible to the Lloyd's network but not the serious ones.

- The Lloyd's reporting network provides a reasonably uniform mechanism that does not introduce any *obvious* biases. There are certainly other networks that are more comprehensive in specific areas but they would introduce considerable bias because of uneven interest in some ships over others - eg. ships that visit American ports or ships insured by Lloyd's or ships of a certain flag etc.

- There is a wide variety of types and severities of incidents reported by Lloyd's. This analysis only considers the fact of the incident, not the type or severity.

The Statistics of Reported Casualties

The purpose of the model is to *quantify the propensity to have casualties* for each of the 2,500 tankers of interest. This can then be used to estimate other things as well, for example, the probability that the ship will become a total loss during the following 12 months.

The method presented is based on the fact that the occurrences of casualties are described very well by conventional probability distributions. The discussion of these distributions is broken up into two parts: First, the number of casualties that occur during any one calender year and second, the number of casualties that have occurred since a tanker first enters service.

Annual Casualty Experience

The first component of the basic model is the use of *the Poisson distribution to describe the number of casualties that a tanker has in one year.* The single parameter of the distribution, λ, is the average number of casualties in a year. This parameter provides the quantification of the propensity to have casualties that we are trying to establish for each tanker. The fleet average is about 0.17 casualties per tanker per year.

An estimate of λ could be made from the tanker's own average annual rate over, for example, the last five years or even over its entire lifetime. This paper describes a different method for estimating λ which takes into account the statistical behavior of the entire fleet.

The second component of the basic model is the use of *the exponential distribution to characterize the variation of the λ's for tankers of a particular age.* That is,

$$\mathrm{pdf}(\lambda \mid \text{age}=a) = e^{-\lambda/\lambda_a}/\lambda_a$$

where λ_a is the average number of casualties for all ships of age 'a'.

The expected value for the probability of 'n' casualties occurring for all of these vessels (as opposed to just one specific vessel) is calculated from:

$$\begin{aligned} E[\,\mathrm{pr}(n \mid \text{age}=a)\,] &= \int_{0\to\infty} \mathrm{pr}(n \mid \lambda) * \mathrm{pdf}(\lambda \mid \text{age}=a)\, d\lambda \\ &= \int_{0\to\infty} e^{-\lambda}\, \lambda^n/n! * e^{-\lambda/\lambda_a}/\lambda_a\, d\lambda \\ &= (\lambda_a)^n \,/\, (1+\lambda_a)^{(n+1)} \end{aligned}$$

showing that the frequency of casualties for ships of the same age are expected to follow a geometric distribution. This formulation can be described as a Bayesian model with a Poisson process, a prior distribution of a degenerate Gamma function (ie. exponential) and a posterior distribution of a degenerate negative binomial (ie. geometric). General derivations are presented by Dropkin.[3]

[This space keeps Figure 1 & its text together on the next page]

Figure 1 shows the actual frequencies of casualties for three different ages. The theoretical results are shown for both the geometric and Poisson distributions with the same λ. The geometric distribution is superior to the Poisson at all three ages. In fact, the geometric distribution matches the actual data well at all ages considered, ie. from 0 to 34.

Figure 1 - *Annual* Casualty Frequencies - Actual Compared to Theory

Table entries are the number of ships that incurred the # of casualties shown in the left hand column. For each age the actual # of ships is shown (Act.), the # predicted by the geometric distribution (Geo.) and the # that would have been predicted had we used the Poisson distribution (Poi.).

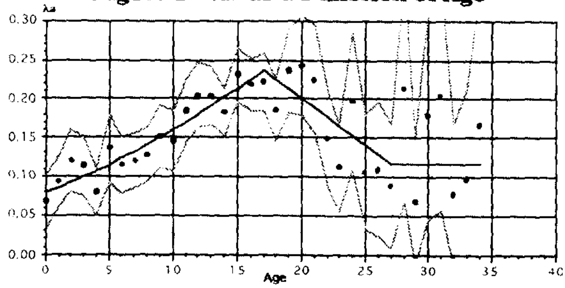
# of	1 Year Old Ships			10 Year Old Ships			15 Year Old Ships		
Casualties	Act.	Geo.	Poi.	Act.	Geo.	Poi.	Act.	Geo.	Poi.
0	353	350.1	348.7	415	412.7	408.7	748	744.8	727.9
1	25	30.1	32.8	48	52.6	59.7	133	139.9	168.1
2	4	2.6	1.5	9	6.7	4.4	31	26.3	19.4
3	1	0.2		1	0.9	0.2	4	4.9	1.5
4					0.1		0	0.9	0.1
5							1	0.2	
Total Ships	383			473			917		
	λ_1 = 0.094			λ_{10} = 0.146			λ_{15} = 0.227		

The ships used at each age to calculate the λ_a's were drawn from all relevant ships at risk over the most recent 5 years. For example, consider the 383 ships that are used to establish λ_1 = 0.094. 89 of these ships were 1 year old (ie. 1 year old at their last 'birthday') on 1/1/92 and they incurred 8 casualties during 1992. Similarly, 77 of the ships were 1 on 1/1/91 and incurred 8 casualties during 1991. 76 ships were at risk during 1990 and had 3 casualties while 74 ships, during 1989, had 6. The oldest group of ships used were 67 vessels that were 1 on 1/1/88 and incurred 11 casualties during 1988.

When all 35 ages (0 to 34) are considered the λ_a's are seen to follow a reasonably smooth progression shown in **Figure 2**. The solid dots are the actual λ_a's calculated from the raw data. The light lines are the limits of the 95% confidence intervals around the actual λ_a's, ie. we are 95% confident that the 'real' λ_a's lie within the band of the light lines. The solid line is just a fitted curve with which age-specfic casualty rates can be conveniently calculated. The confidence intervals are determined from the variance of the casualty ratio which, for the geometric distribution, is given by $\lambda_a*(1+\lambda_a)/\#$ ships.

At this time, there is no completely satisfactory explanation for the drop in casualty rate after 17 years.

Figure 2 - λ_a as a Function of Age



Lifetime Casualty Experience

Since the annual number of casualties, n_{ta}, of tanker 't' at age 'a', follows a Poisson distribution, with parameter λ_{ta}, then the number of casualties accumulated by a single tanker after 'a' years, N_{ta}, must also follow a Poisson distribution, with parameter Λ_{ta} given by:

$$\Lambda_{ta} = \Sigma_k \lambda_{tk}, \qquad k = 0 \text{ to } a-1$$

Capitals indicate *lifetime*, or cumulative, (as opposed to *annual*) variables or parameters.

That the lifetime experience follows a Poisson distribution is demonstrated through iterative convolutions of the annual experience. In general,

$$PR(N \mid yrs=a) = \sum_n PR(N-n \mid yrs=a-1) * pr(n \mid age=a-1), \qquad n = 0 \text{ to } N$$

where 'PR' indicates the probability for the lifetime number of casualties and 'pr', the probability for the annual number.[4] At the end of the second year, for example, this becomes

$$\begin{aligned} PR(N \mid yrs=2) &= \sum_n PR(N-n \mid yrs=1) * pr(n \mid age=1), && n = 0 \text{ to } N \\ &= \sum_n pr(N-n \mid age=0) * pr(n \mid age=1), && n = 0 \text{ to } N \\ &= \sum_n e^{-\lambda t0} \lambda_{t0}^{(N-n)}/(N-n)! * e^{-\lambda t1} \lambda_{t1}^{n}/n!, && n = 0 \text{ to } N \\ &= e^{-(\lambda t0+\lambda t1)} (\lambda_{t0}+\lambda_{t1})^N/N! \end{aligned}$$

ie, a Poisson distribution with parameter $\Lambda_{t2} = \lambda_{t0} + \lambda_{t1}$. Repeated convolutions yield Poisson parameters, Λ_{ta}, given by:

$$\begin{aligned} \Lambda_{ta} &= \Lambda_{t(a-1)} + \lambda_{t(a-1)}, && \text{where } \Lambda_0 = 0 \\ &= \Sigma_k \lambda_{tk}, && k = 0 \text{ to } a-1 \end{aligned}$$

Because the λ_{ta}'s vary with age it is not clear how the Λ_{ta}'s ought to vary across the fleet for any given age. This is because the distribution of the sum of independent variables, such as the λ_{ta}'s, even with simple distributions, like the exponential, are usually difficult. It turns out, in this case though, that the Λ_{ta}'s, like the λ_{ta}'s, are also distributed exponentially. This is implied from the fact that the frequency of *lifetime* casualties, like the frequency of *annual* casualties, nearly follows a geometric distribution. That the Λ_{ta}'s are distributed exponentially is crucial to the basic model and discussed further in the section "Calculating the Expected Casualty Rate".

Figure 3 shows the actual lifetime frequencies of casualties for three different ages. The theoretical results are shown for both the geometric and Poisson distributions with the same Λ. The geometric distribution matches the actual data reasonably well at all three ages, while the Poisson grossly deviates at the larger values of Λ for 10 and 15 year old tankers. The geometric distribution matches the actual data reasonably well at all ages considered, ie. from 0 to 34.

Figure 3 - *Lifetime* Casualty Frequencies - Actual Compared to Theory

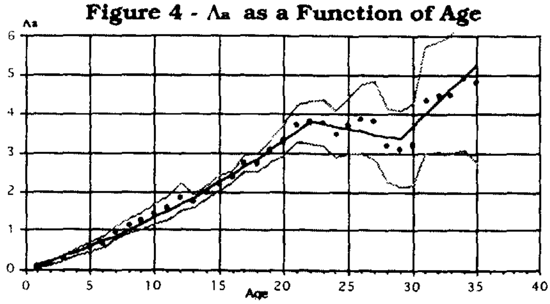
Table entries are the number of ships that incurred the # of casualties shown in the left hand column. For each age the actual # of ships is shown (Act.), the # predicted by the geometric distribution (Geo.) and the # that would have been predicted had we used the Poisson distribution (Poi.).

# of	1 Year Old Ships			10 Year Old Ships			15 Year Old Ships		
Casualties	Act.	Geo.	Poi.	Act.	Geo.	Poi.	Act.	Geo.	Poi.
0	335	331.1	326.4	184	182.9	96.5	238	268.7	82.4
1	39	44.8	52.2	114	112.2	153.3	238	189.9	198.5
2	7	6.1	4.2	63	68.8	121.9	137	134.3	239.2
3	1	0.8	0.2	48	42.2	64.6	104	95.0	192.1
4	1	0.1		24	25.9	25.7	67	67.1	115.8
5				12	15.9	8.2	35	47.5	55.8
6				8	9.7	2.2	17	33.6	22.4
7				10	6.0	0.5	15	23.7	7.7
8				4	3.7	0.1	15	16.8	2.3
≥9				6	5.7		51	40.4	0.8
Total Ships	383			473			917		
	Λ_1= 0.16			Λ_{10}= 1.59			Λ_{15}= 2.41		

There is a slight systematic difference between the actual frequencies and those given by the geometric distribution for tankers that have been in service longer than 10 years. The number of ships with no casualties is overstated while the number of ships with one casualty is understated. This effect, seen in **Figure 3** for the 15 year old ships, is discussed later in the section "Modifications to the Basic Model".

When all 35 ages (0 to 34) are considered the Λ_a's are seen to follow a reasonably smooth progression shown in **Figure 4**. The solid dots are the actual Λ_a's calculated from the raw data. The light lines are the limits of the 95% confidence intervals around the actual Λ_a's, ie. we are 95% confident that the 'real' Λ_a's lie within the band of the light lines. The solid line is just a fitted curve with which age-specfic lifetime casualty rates can be conveniently calculated. The confidence intervals are determined from the variance of the lifetime casualty ratio given

Figure 4 - Λ_a as a Function of Age



by $\Lambda_a*(1+\Lambda_a)/\#$ ships. The drop in casualty rate between 22 and 29 years is not satisfactorily explained at this time.

Calculating the Expected Casualty Rate

The calculation of λ_{ta}, the casualty rate for tanker 't' at age 'a', is based on the assumption that $\lambda_{ta}/\lambda_a = \Lambda_{ta}/\Lambda_a$.

This assumption follows from an effort to understand why the Λ_{ta}'s are exponential. An explanation could be that the λ_{ta}'s are not really independent at all because λ_{ta}/λ_a remains more or less constant over a tanker's lifetime. This condition eliminates the complexities of convolutions and assures that the lifetime casualty rates will be exponential. It also implies that Λ_{ta}/Λ_a will be constant and have the same value, hence $\lambda_{ta}/\lambda_a = \Lambda_{ta}/\Lambda_a$ as specified.

Special importance is assigned to this ratio because of its persistence. It will be referred to as the 'casualty relativity', 'R_t', of the ship because it specifies an individual ship's risk relative to the rest of the fleet.

The calculations proceed in three steps: Bayes' theorem is first used to calculate an expected value for Λ_{ta} given N_{ta} as described below. Then, R_t is calculated from $R_t = \Lambda_{ta}/\Lambda_a$. Finally, the expected value of λ_{ta} is calculated from $\lambda_{ta} = R_t * \lambda_a$.

The development of $E[\Lambda_{ta}]$ begins with

$$E[\Lambda_{ta} \mid N_{ta}] = \int_{0\to\infty} \Lambda * \mathrm{pdf}(\Lambda \mid \mathrm{age}=a, N_{ta})\, d\Lambda,$$

where $\mathrm{pdf}(\Lambda \mid \mathrm{age}=a, N_{ta})$ is obtained from Bayes' theorem as follows

$$\mathrm{pdf}(\Lambda \mid \mathrm{age}=a, N_{ta}) = \frac{\mathrm{pf}(N_{ta} \mid \Lambda) * \mathrm{pdf}(\Lambda \mid \mathrm{age}=a)}{\mathrm{pf}(N_{ta})} = \frac{e^{-\Lambda}\, \Lambda^{N_{ta}}/N_{ta}! * e^{-\Lambda/\Lambda_a}/\Lambda_a}{(\Lambda_a)^{N_{ta}} / (1+\Lambda_a)^{(N_{ta}+1)}}$$

Hence,

$$E[\Lambda_{ta} \mid N_{ta}] = \frac{(1+N_{ta}) * (\Lambda_a)^{(N_{ta}+1)} / (1+\Lambda_a)^{(N_{ta}+2)}}{(\Lambda_a)^{N_{ta}} / (1+\Lambda_a)^{(N_{ta}+1)}} = \frac{(1+N_{ta})}{(1+\Lambda_a)} * \Lambda_a,$$

which yields $R_t = (1+N_{ta}) / (1+\Lambda_a)$. λ_{ta} is then calculated as $R_t * \lambda_a$.

Modifications to the Basic Model

There are two modifications that are made to the basic model as described above.

The first relates to the *systematic error in lifetime casualties for ships that have been in service more than 10 years*. The basic model overstates the number of these ships which have not incurred any casualties and correspondingly understates the number that have incurred only one. Other frequencies are predicted accurately.

A change in the assumed distribution of Λ's from the exponential to the more general translated Gamma function eliminates this systematic error. The effect of this change on the calculation of R was studied for all ages between 9 and 23 for N = 0,1,2 and 3. It was found that for N = 0 the basic model predicted a smooth drop in R from 0.42 at 9 years old down to 0.22 at 23 years whereas the more accurate model yielded a constant R of 0.42 between 9 and 23 years of service. Similarly for N = 1 the more accurate model predicts a more or less constant R of 0.50 for ships older than 13 years. For N = 2 and 3 there was no significant difference between the basic model and the more accurate model.

The basic model has been modified by replacing the R value calculated by the exponential model by the constant value found above. This method of making the modification was chosen for two reason: First, the calculations with the translated Gamma function are much more complex and time consuming than those with the exponential hence avoiding them with no loss in accuracy is convenient. Second, the roughly 800 ships that are affected by this are at *below*-average risk whereas the value of the model is in its ability to accurately quantify the risk of those ships that are at *above*-average risk.

The second modification results from *the basic model's tendency to exaggerate the deviation of a tanker's casualty rate from the average rate*. For example, tankers that the model identifies as being at high risk, do have many casualties, but not quite as many as predicted. Similarly, tankers identified as being at low risk, do have very few casualties, but slightly more than predicted.

At this time there is no satisfactory explanation for this 'regression towards the average' but, nonetheless, a satisfactory, heuristic correction is made with:

$$R_{corrected} = R^{0.74}$$

For all practical purposes the range of corrected R-values is 1/2 to 3.

Comparison of Expected Casualty Rates with Actual Rates

Casualties in 1991

On 1/1/91 there were 2,420 privately-owned tankers. The basic model was applied using the same kind of casualty information presented earlier but only using data before 12/31/90. Values for λ were calculated for all of the ships. Adding up all the λ's yielded a total prediction of 436.6 casualties for an average rate of 0.180 (436.6/2420). The fleet was then broken up into the 6 risk groups shown in **Figure 5**.

Figure 5 - 1991 Casualties, Pred. & Act.

Risk Group	Casualty Rate Range	#Ships	# of Casualties — Predicted	Actual	z
0.5	-0.135	1043	105.0 ±10.7	101	-0.37
1.0	0.135-0.225	820	139.4 ±12.8	124	-1.20
1.5	0.225-0.315	299	79.9 ±10.1	94	+1.40
2.0	0.315-0.405	128	45.1 ± 7.8	39	-0.78
2.5	0.405-0.495	72	32.0 ± 6.8	40	+1.18
3.0	0.495-	58	35.2 ± 7.5	27	-1.09
		2420	436.6 ±22.7	425	

The lowest risk group includes ships with λ's less than 75% of the average rate. These 1043 ships (43%) collectively had an actual casualty rate of 0.097 (101/1043) while the predicted rate was 0.101 (105.0/1043). Since this rate is about half the fleet average of 0.180, the group is labeled '0.5'.

The average risk group, labeled '1.0', consists of all ships with λ between 75% and 125% of the average rate. With 820 ships, this group is 34% of the fleet. Collectively they experienced a casualty rate of 0.151 (124/820) while the predicted rate was 0.170 (139.4/820).

The remaining 557 (23%) of the ships are spread between the four high risk groups which run from 1.5 up to 3 times the average rate. These ships, all taken together, had a casualty rate of 0.359 (94+39+40+27/557), twice the average rate, while the predicted rate was 0.345 (79.9+45.1+32.0+35.2/557).

The predictions seem to match the actual results well but verifying this requires that the difference between the actual and predicted number of casualties be looked at carefully. These differences are expected to be the result of the Poisson processes themselves and not 'error'. In this sense, these differences are part of the prediction - they *must* occur, otherwise the model cannot be correct. The issue, then, is determining whether the actual differences are consistent with the statistics of the model. To do this, the z-values given by $z = (A-P)/\sigma$ are considered - 'A' and 'P' are the predicted and actual number of casualties and σ is the expected standard deviation. Taken all together, the z-values should behave like a random sample from the unit normal distribution, N(0,1).

The mean of the z's is 0.14 ($p>.8$) with $\sigma = 1.15$ ($p>.3$). These values are comfortably consistent with N(0,1). Further, there is no evidence of skewness (coef. of skew= -0.52) and only slight evidence of negative kurtosis (coef. of kur= -1.50). There were no tables available to calculate p values for skew and kurtosis since n = 6 is so small. An alternate measure, a 3 df Chi2 test constructed to maximize the effect of any kurtosis, yielded p=0.29.

Casualties in 1992

On 1/1/92 there were 2,507 privately-owned tankers. The basic model was again applied using, in addition, the casualties incurred during 1991. New values for λ were calculated for all of the ships. Adding up all the λ's yielded a total prediction of 448.8 casualties for an average rate of 0.179 (448.8/2507). **Figure 6** shows the fleet broken up into the same 6 *relative* risk groups.

The first thing to note for 1992 is that the actual number of reported casualties, 341, is only 76% of the total predicted. This is 4.7 σ's below the prediction ($p<0.0002\%$). This large difference is unlikely to be part of normal variation. Possibly this reduction is due to the increased deductibles and exclusions mentioned in the introduction, resulting in more casualties going unreported.

Figure 6 - 1992 Casualties, Pred. & Act.

Risk Group	Casualty Rate Range	#Ships	# of Casualties Predicted	Actual
0.5	-0.134	1115	111.9 ±11.1	95
1.0	0.134-0.224	812	135.8 ±12.6	98
1.5	0.224-0.313	313	83.1 ±10.2	67
2.0	0.313-0.403	137	48.4 ± 8.1	32
2.5	0.403-0.492	59	25.8 ± 6.1	21
3.0	0.492-	71	43.8 ± 8.4	28
		2507	448.8 ±23.0	341

If all of the λ's are scaled down to 76% of their calculated value we will still be able to assess the model's ability to quantitatively discriminate between the different risk levels. **Figure 7** shows that the scaled-down predictions agree well with the actuals.

Figure 7 - 1992 Casualties, Pred. & Act.
Revised Predictions

Risk Group	Casualty Rate Range	#Ships	# of Casualties Revised Pred.	Actual	z
0.5	-0.102	1115	85.0 ± 9.6	95	+1.04
1.0	0.102-0.170	812	103.2 ±10.8	98	-0.48
1.5	0.170-0.238	313	63.1 ± 8.7	67	+0.45
2.0	0.238-0.306	137	36.8 ± 6.8	32	-0.71
2.5	0.306-0.374	59	19.6 ± 5.1	21	+0.27
3.0	0.374-	71	33.3 ± 7.0	28	-0.76
		2507	341.0	341	

The z-values have a mean of -0.03 and σ = 0.73, again, comfortably consistent with the expected N(0,1). The skewness of 0.33 continues to be insignificant and the kurtosis is -1.26. The two year persistence of negative kurtosis is noted with no explanation.

Predicting the Risk of Total Loss

There is a direct relation between the probability of a ship becoming a total loss and its predicted casualty rate. This is established by analyzing all 202 total losses, both actual and constructive, that have occurred to privately-owned tankers since 1976. The basic model was applied to each of these ships based on their age and casualty record on 1/1 of the year they were lost. Account was made, in these calculations, for the fact that more casualties were reported in the '60s and early '70s than are reported now.

The TLs were then grouped into the same 6 risk groups used before. **Figure 8** shows an estimate of how many ship-years at risk there have been, in each risk group, during the 17 years since 1976. These estimates were made by: First, assuming 2,500 ships were at risk each year. Second, assuming the percentage of ships in each risk group has been relatively constant and can be estimated by averaging the percentages in the '91, '92 and '93 fleets. The ratio of TLs to number of ships at risk is then given along with its σ. The σ shown reflects no model error, only variation implied by the statistics of the Poisson distribution.

The three high risk groups are pooled in the last line of **Figure 8**. The ships in this pool, with risk = 2.4 (ie. casualty rate is 2.4 times the average), *are 3x more likely to be reported as total losses than average and 10x more likely to be reported as total losses than the low risk ships.*

There is no formal prediction of total loss rates to compare with the actual values, however, it is seen that, generally, as the risk goes up, the rate of TLs goes up. A straight line of total loss rate vs risk fits well to the three lowest risk groups and the pooled high risk group. The line is given by 0.69*(risk-0.31) and can be used as a 'predictor' for the total loss probability of a ship where risk = $\lambda_t/\lambda_{average}$.

Figure 8 - 202 TLs from '76 to '92
Actual and Fitted

Risk Group	#Ships	#TLs	TL rate	Fit
0.5	18,600	24	0.13% ±0.03	0.13%
1.0	14,100	74	0.52% ±0.06	0.48%
1.5	5,300	37	0.70% ±0.11	0.82%
2.0	2,300	40	1.74% ±0.27	1.17%
2.5	1,100	18	1.64% ±0.39	1.51%
3.0	1,100	9	0.82% ±0.27	1.86%
	42,500	202	0.48% ±0.03	0.48%
Pool of 3 High Risk Groups				
2.4	4,500	67	1.49% ±0.18	1.44%

In actual practice, the total loss probabilities for all ships are scaled, after being calculated, so that 7 total losses are predicted for the coming year since this has been the consistent fleet experience since 1985.

Summary and Areas of Further Research

The statistics of oil tanker casualties reported in Lloyd's List are found to follow Poisson's distribution for individual tankers while the Poisson parameters for all tankers of the same age are found to follow exponential distributions. Bayes' theorem permits the calculation of the casualty relativity, R, for each ship given its age and lifetime number of casualties. An estimate of a ship's casualty rate is made by multiplying the average casualty rate for tankers of the same age by R.

The predicted casualty rates permit the tankers to be separated into six risk groups in order to check their accuracy. The predicted number of casualties for each group was found to be consistent with the number actually incurred

The usefulness of the predicted casualty rate was demonstrated by showing that the probability of total loss correlates with the predicted rate.

Oil spills and other serious casualties are currently being examined as their frequency varies with λ and also with age and R separately.

There are three other areas which need additional attention:

The uncertainty in λ

Experience-based ratings can suffer from the infrequency of the events. The principal effect that this has on the λ's is a relatively large variation for a given N. [σ for λ is $(1+N)^{1/2}/(1+\Lambda)$ as compared to the expected value of $(1+N)/(1+\Lambda)$.] One purpose of Dropkin's paper, in fact, was to point out this problem in the arena of auto insurance for individuals. For oil tanker casualties, it could be useful to address this problem by using additional information to select a value for λ slightly different from its expected value. For example, a tanker with an owner who has few casualties, could be assigned a λ somewhat less than the expected value, while a tanker with an owner who has many casualties could be assigned a λ somewhat greater than the expected value.

R may change with time

The basic model assumes that the casualty relativity, R, is constant throughout a tanker's lifetime. There are some circumstances, though, where this may not be reasonable, for example, after a tanker is sold to a new owner. It would be desirable to identify, as quickly as possible, when recent casualty experience may indicate a change from the historical experience.

Utilizing claims information

Establishing a relationship between actual claims and λ could increase the utility of the model.

Notes and References

1 Flitner and Brunck, *Ocean Marine Insurance*, Insurance Institute of America discusses all aspects of marine insurance in detail.

2 Borch, Karl, "Mathematical Models for Marine Insurance," *Scandinavian Actuarial Journal*, 1979, pp 25-36 discusses factors other than risk that enter into marine insurance.

3 Dropkin, Lester, "Some Considerations on Automobile Rating Systems Utilizing Individual Driving Records", *Proceedings of the Casualty Actuarial Society* XLVI, 1959, pp 165-176.

4 Strictly speaking 'n' in these equations is 'n_{ta}' and 'N' is '$N_{t(a-1)}$'. The 't' and age designating subscripts are dropped to improve readability.