Entropy under partial insurance

Athanasios Sachlas and Takis Papaioannou

Department of Statistics & Insurance Science, University of Piraeus, 185 34, Piraeus, Greece

Abstract

Often insurers construct insurance policies where the claim paid to the client is part of the loss that occurs. This is the result of using deductibles and policy limits. In this paper we explore the uncertainty on the losses under partial insurance coverage. The measure of uncertainty we use is Shannon's entropy [17], which is defined as $H(X) = -\sum_i p_i \ln p_i$ or $H(X) = -\int f(x) \ln f(x)\, dx$, in the case of discrete or continuous distributions, respectively. It is a useful notion in Statistics, as it is a measure of the uncertainty related to a random variable X. It is a descriptive measure of distributions belonging to the class of dispersion measures, such as the variance and the standard deviation. Analytic expressions for the entropy of the most known loss models for several kinds of partial insurance coverage are presented.

Keywords. entropy, loss distributions, partial insurance, policy limits, deductibles

1 Introduction

In many insurance policies the amount paid by the insurer is not necessarily the total amount of the claim but it is part of the loss that occurs. There are several standard types of partial insurance coverage such as per payment insurance, excess of loss insurance, residual insurance etc.

The best policy for an insurance company is that which lasts for a long period and it is less uncertain with reference to its claims. A well known measure of uncertainty associated with a random variable, comes from the field of information theory, and is called *entropy*, *Shannon's entropy*, and is given by

$$H(X) = -\sum_x p(x) \ln p(x) \quad \text{or} \quad H(X) = -\int f(x) \ln f(x)\, dx$$

depending on whether the random variable X is discrete or continuous, respectively [17]. In the latter case, $H(X)$ is also called *differential entropy*. In the sequel $H(X)$ will be denoted by H_X. Entropy quantifies the expected uncertainty related with the result of an experiment. In other words it provides information for the predictability of the result of a random variable X. The bigger the entropy the less concentrated the distribution of X and thus an observation of X provides a little information. Shannon's entropy, along with other measures of entropy such as the Rényi entropy, may be regarded as descriptive quantity of the corresponding pdf. The entropy serves as a measure of variability for continuous variables or as a measure of variation or diversity of the possible values of a discrete variable ([10], [13], [18]). Other measures of entropy, and generalizations thereof, have been proposed in the literature (c.f. [3], [15]). Due to the widespread applicability and use of entropy, the derivation of explicit expressions for the Shannon entropy for univariate and multivariate distributions has been a subject of interest. Several authors have dealt with this matter, among them [2], [9] and more recently [5], [13], [14] and [19].

Highly uncertain insurances are less reliable. The uncertainty for the loss in an insurance policy can be quantified by the entropy of the corresponding loss distribution. However, frequently in actuarial practice, the practitioner has at hand transformed data as a consequence of deductibles and policy (liability) limits.

The purpose of this paper is to investigate the uncertainty under several partial insurance schemes and to provide explicit expressions of Shannon's entropy for the most known loss distributions under these situations that frequently appear in actuarial practice. It complements previous work of the authors on the entropy under inflation and truncation from above [16]. The entropy of exponential, Pareto, Weibull, Gamma, Burr and Generalized Pareto distributions are presented among other in [16, Table1]. In Section 2 we investigate the effect on entropy of the use of a deductible while the effect of policy limits is investigated in Section 3. Analytical expressions are given for each of the above mentioned situations. The entropy of the classical ruin problem is presented in Section 4 while concluding remarks are given in Section 5. For reasons of brevity proofs of our results are omitted or hinted.

2 The effect of truncation from below - losses with a deductible

Suppose that losses are not recorded or reported below a specified amount. In this case the data are truncated from below or left truncated. The most common reason of left truncation is the use of a deductible. Let X be the absolutely continuous random variable of losses with pdf $f(x)$ and cdf $F(X)$, $x > 0$ and d the deductible value. Following the terminology of Kluggman, Panjer and Willmot [12], the random variable expressing losses associated with an insurance

policy having a deductible d is given in two ways:
(i) *per payment*

$$Y_p(d) = \begin{cases} X, & X > d \\ \text{not defined}, & X \leq d. \end{cases} \quad \text{equivalently} \quad Y_p(d) = X | X \geq d$$

(ii) *per loss* (franchise deductible)

$$Y_l(d) = \begin{cases} X, & X > d \\ 0, & X \leq d. \end{cases}$$

The cost per loss in the latter case concludes the zero amounts paid by the insurer for losses not exceeding the deductible d. The difference between these two definitions is that the latter assigns a positive probability mass at the zero point (when $X \leq d$) thus making $Y_l(d)$ a not absolutely continuous random variable.

Let us now investigate the effect of truncation from below. Differential entropy under this situation is a useful measure as it measures the uncertainty about the loss covered by the insurance company. The relations between the entropies of X, $Y_p(d)$ and $Y_l(d)$ are given by the lemmas below.

Lemma 1 *The per payment entropy of the random variable $Y_p(d)$ of losses truncated from below or the per payment losses under a deductible, is given by*

$$H(f,d) = \frac{1}{\bar{F}_X(d)} \left[H_X + \int_0^d f_X(x) \ln f_X(x)\, dx + \bar{F}_X(d) \ln \bar{F}_X(d) \right],$$

where H_X is the entropy of distribution X, $\bar{F}_X(x) = 1 - F_X(x)$, where $F_X(x)$ is the cumulative distribution function of X and d is the deductible.

The next lemma gives the per loss entropy expression.

Lemma 2 *The per loss entropy of the random variable $Y_l(d)$ of losses truncated from below, is given by*

$$H_l(f,d) = H_X + \int_0^d f_X(x) \ln f_X(x)\, dx - F_X(d) \ln F_X(d),$$

where H_X is the entropy of distribution X, $F_X(x)$ is the cumulative distribution function of X and d is the deductible.

Ebrahimi in [7] presented the following expression of $H(f,d)$ in terms of the hazard function $\lambda_X(x)$ of X,

$$H(f,d) = 1 - \frac{1}{\bar{F}_X(d)} \int_d^\infty (\log \lambda_X(x))\, f_X(x)\, dx$$

Table 1: Per payment entropy

Distribution	**Entropy** $H(f,d)$
Exponential	$1-\ln\lambda$
Pareto	$1+\frac{1}{a}-\ln\alpha+(1+2\alpha)\ln(\lambda+d)$
Weibull	$1-\ln c-\ln\tau-(\tau-1)\ln d-\frac{(\tau-1)e^{cd^\tau}\Gamma(0,cd^\tau)}{\tau}$
Gamma	$\frac{1}{\varphi(\alpha,\lambda d)}\Big\{\Gamma(\alpha+1)-\Gamma(\alpha+1,\lambda d)$ $+(\alpha-1)\Gamma(\alpha,\lambda d)\ln d+(\alpha-1)G^{3,0}_{2,0}\left(\lambda d\,\middle\vert\,\begin{matrix}\mathbf{a}\\ \mathbf{b}\end{matrix}\right)$ $-\Gamma(\alpha)\left[\alpha+(\alpha-2)\ln\lambda+\ln\xi(\alpha,\lambda d)+\Gamma(\alpha,\lambda d)\ln\frac{\lambda^\alpha}{\xi(\alpha,\lambda d)}\right]\Big\}$
Burr	not in a closed form
Gen. Pareto	not in a closed form

$\varphi(\alpha,x)=\Gamma(\alpha)\left(\Gamma(\alpha,x)-1\right)$ $\quad \mathbf{a}=(1,1)$ $\quad \Gamma(\alpha,x)=\int_x^\infty t^{\alpha-1}e^{-t}\,dt$
$\xi(\alpha,x)=\Gamma(\alpha)\left(1-\Gamma(\alpha,x)\right)$ $\quad \mathbf{b}=(0,0,\alpha)$

as a measure of uncertainty of the residual lifetime distribution, i.e. given that an item has survived up to time t and is the dynamic version of the classical entropy. Analogous measures of information for multivariate distributions when their supports are truncated progressively are proposed by [8]. No analytic expressions have been presented in bibliography, apart from [4] for the exponential and the Pareto distribution.

Obviously $H_l(f,d)$ cannot be expressed in terms of hazard since $Y_l(d)$ is not absolutely continuous.

Analytic expressions for the per payment and per loss entropy are shown in Tables 1 and 2, respectively. Truncation from above does not affect the entropy of the exponential distribution. In the entropy expressions for the Gamma distribution the Meijer function

$$G^{p,q}_{m,n}\left(z\,\middle\vert\,\begin{matrix}\mathbf{w}\\ \mathbf{x}\end{matrix}\right)=\frac{1}{2\pi i}\int_L\frac{\prod_{j=1}^m\Gamma(x_j-s)\prod_{j=m+1}^q\Gamma(1-w_j+s)}{\prod_{j=1}^n\Gamma(1-x_j+s)\prod_{n=1}^p\Gamma(w_j-s)}z^s\,ds,$$

$\mathbf{w}=(w_1,\ldots,w_q)^T$, $\mathbf{x}=(x_1,\ldots,x_q)^T$, is involved [1]. For some models, such as the Burr distribution and the Generalized Pareto distribution the Shannon entropy has no closed form, so a numerical method should be employed in order to compute their value.

The above presented complicated formulae do not permit us to give a general and straightforward interpretation of the effect of the deductible for the various loss models. However we can see that truncation from below does not affect the entropy of the exponential distribution.

The two entropies are related with the following relationship

$$H_l(f,d)=\bar{F}_X(d)H(f,d)-F_X(d)\ln F_X(d)-\bar{F}_X(d)\ln\bar{F}_X(d),$$

Table 2: Per loss entropy

Distribution	**Entropy** $H_l(f,d)$
Exponential	$e^{-\lambda d}(1+\lambda d-\ln\lambda)-(1-e^{-\lambda d})\ln(1-e^{-\lambda d})$
Pareto	$1+\frac{1}{\alpha}+\ln\lambda-\left(\frac{\lambda}{\lambda+d}\right)^{\alpha}\ln\alpha-\left[1-\left(\frac{\lambda}{\lambda+d}\right)^{\alpha}\right]\left(\alpha\ln\lambda-\ln\left(\frac{\lambda}{\lambda+d}\right)^{\alpha}\right)$
Weibull	$-\ln\tau+(\tau-1)\frac{\gamma+\ln c}{\tau}+(1+cd^{\tau})e^{-cd^{\tau}}$ $-\frac{\tau-1}{\tau}\left(\gamma+\Gamma(0,cd^{\tau})+e^{-cd^{\tau}}\tau\ln d\right)+(1-e^{-cd^{\tau}})\ln\frac{c\tau}{1-e^{-cd^{\tau}}}$
Gamma	$\alpha+\ln\Gamma(\alpha)-(\alpha-1)\psi(\alpha)-\ln\lambda-\frac{\Gamma(\alpha+1)-\Gamma(\alpha+1,\lambda d)}{\Gamma(\alpha)}$ $-\frac{\alpha-1}{\Gamma(\alpha)}\left[\Gamma(\alpha,\lambda d)\ln d+G^{3,0}_{2,0}\left(\lambda d\,\middle\vert\,\begin{matrix}\mathbf{a}\\ \mathbf{b}\end{matrix}\right)-\Gamma(\alpha)(\psi(\alpha)+\ln\lambda)\right]$ $+\Gamma(\alpha,\lambda d)\left(\alpha\ln\lambda-\ln\Gamma(\alpha)-\ln\Gamma(\alpha,\lambda d)\right)$
Burr	not in a closed form
Gen. Pareto	not in a closed form

$\mathbf{a}=(1,1)\quad \mathbf{b}=(0,0,\alpha)\quad \Gamma(\alpha,x)=\int_x^{\infty}t^{\alpha-1}e^{-t}\,dt$

Proposition 1 *If $H_X<0$ then $H(f,d)<H_l(f,d)$ for all d while if $H_X>0$ then it is inconclusive.*

Related to the previous losses with a deductible are the residual (after d) or excess losses defined by $X-d$:
(i) *residual loss*

$$Z_p(d)=\begin{cases}X-d, & X>d\\ \text{not defined}, & X\le d\end{cases}\quad \text{equivalently}\quad Z_p(d)=X-d|X>d$$

(ii) *excess loss*

$$Z_l(d)=\begin{cases}X-d, & X>d\\ 0, & X\le d.\end{cases}$$

In these situations the insurer pays the amount of loss in excess of d if the loss is greater than d. In the cost per loss of the latter scheme the zero amounts paid by the insurer for losses smaller than d are concluded. $Z_l(d)$ is not an absolutely continuous distribution. The following lemma gives the expressions for the per payment residual loss and per loss residual entropy.

Lemma 3 *(i) The per payment residual loss entropy of $Z_p(d)$ is identical with the (simple) per payment entropy.*
(ii) The per loss residual entropy of $Z_l(d)$ is identical with the (simple) per loss entropy.

These results are natural due to shift invariant property of Shannon's entropy.

We then present some properties of the per payment entropy $H(f,d)$.

Proposition 2 *$H(f,d)$ is independent of d if and only if $\lambda_X(x)$ is constant.*

In other words, if the risk is constant, uncertainty with a deductible d does not depend on the deductible.

Proposition 3 *If X is absolutely continuous and $H(f,d)$ is increasing in d then $H(f,d)$ uniquely determines $F_X(x)$.*

Proposition 4 *The entropy of residual losses is bounded by the mean residual losses*

$$\delta_p(d) = E(X-d|X>d) = E(X|X>d) - d = \frac{\int_d^\infty \bar{F}_X(x)\,dx}{\bar{F}_X(x)} - d.$$

More specifically it holds

$$H(f,d) \leq 1 + \ln \delta_p(d).$$

Proposition 5 *The entropy $H(f,d)$ of truncated from below losses (i) is increasing in d if $\lambda_X(x)$ is decreasing in x and (ii) is decreasing in d if $\lambda_X(x)$ is increasing in x.*

Proposition 6 *If the mean residual loss is decreasing in d then the uncertainty $H(f,d)$ is also decreasing in d.*

Proposition 7 *If $H(f,d)$ is increasing in d then its mean residual loss is also increasing.*

Proposition 8 *Consider two insurance policies with risk functions $\lambda_1(x)$ and $\lambda_2(x)$ such that $\lambda_2(x) = h(x)\lambda_1(x)$. Let also $h(x)$ be increasing in x and $0 < h(x) < 1$. Then if Policy 1 has a decreasing mean residual loss then Policy 2 has also a decreasing mean residual loss.*

3 The effect of truncation from above

The aim of this section is to study how truncation from above affects the entropy. Suppose that losses X are not recorded for or above a certain liability limit u. In this case the data W are truncated from above or, in other words, right truncated. In this case, the random variable modeling the truncated from above losses is

$$W(u) = \begin{cases} X, & X \leq u \\ \text{not defined}, & X > u. \end{cases}$$

Equivalently $W(u) = X|X \leq u$. For example, a car insurance policy covers losses up to a limit u, while major losses are covered by the car owner. If the loss is X then the loss for the insurance company is $X|X \leq u$. We note that the loss with truncation from above is different from the loss with censoring from above, which is defined as the random variable $V(u) = min\{X, u\}$. In this case, if the loss is $X > u$, the insurance company pays an amount u. This corresponds to the per loss case of Section 2.

The relationship between the entropies of X and $W(u)$, is given by the following lemma (cf [16]).

Lemma 4 *The entropy of the random variable $W(u)$ of observed truncated from above data, is given by*

$$\bar{H}(f,u) = \frac{1}{F_X(u)}\left[H_X + \int_u^\infty f_X(x)\ln f_X(x)\,dx + F_X(u)\ln F_X(u)\right],$$

where H_X is the entropy of X, $F_X(x)$ is the distribution function of X and u is the point of truncation.

Proposition 9 *$\bar{H}(f,u) < \ln u$ for all $u > 0$.*

Di Crescenzo and Longobardi in [6] presented the following expression of $\bar{H}(f,u)$ in terms of the reverse hazard function $\tau_X(x) = f(x)/F(x)$,

$$\bar{H}(f,u) = 1 - \frac{1}{F(u)}\int_0^u f(x)\ln \tau_X(x)\,dx$$

as an entropy-based measure of uncertainty in past lifetime distributions, called past lifetime entropy. No analytic expressions have been presented in the bibliography.

The relevant analytic expressions of Shannon's entropy $\bar{H}(f,u)$ are given in [16, Table 2]. We note that for the above mentioned distributions, $F_X(\cdot)$ is explicitly known. The Meijer G function, is aslo involved in the Gamma distribution's entropy enxpression. This makes it difficult to give a general and straightforward interpretation of the effect of truncation from above on the entropy.

The next proposition establishes a relationship between H_X, $H(f,d)$ and $\bar{H}(f,d)$.

Proposition 10 *For all $d > 0$,*

$$H_X = H\left[F_X(d), \bar{F}_X(d)\right] + \bar{F}_X(d)H(f,d) + F_X(d)\bar{H}(f,d)$$

where $H[p,\bar{p}] = -p\ln p - \bar{p}\ln\bar{p}$ is the entropy of a Bernoulli distribution.

Proposition 10, assuming that the deductible and the liability limit have the same value, shows that the uncertainty of a loss is decomposed into three parts: (i) the uncertainty of whether the loss has exceeded or not the deductible d, (ii) the uncertainty about the loss having exceeded the value of deductible d, i.e. the loss covered by the insurance company, and (iii) the uncertainty about the loss, covered by the insurance company, given that the loss is smaller than the liability limit d.

Remark 1 *Entropy under truncation from above can also be computed by the expression*

$$\bar{H}(f,u) = \ln F_X(u) - \frac{1}{F_X(u)} \int_0^u f_X(x) \ln f_X(x)\,dx.$$

Two other interesting quantities in loss distributions and actuarial mathematics are $W_1(u) = u - W$, which describes the "remaining" of the coverage and $W_2(u) = X - u$, which describes the "remaining" of the claim. Because differential entropy is invariant under shifts we have:

$$H_{W_1(u)} = H_{W(u)} \quad \text{and} \quad H_{W_2(u)} = H_X.$$

We have already mentioned that the loss with truncation from above is different from the loss with censoring from above, which is defined as the random variable

$$V(u) = \begin{cases} X, & X \leq u \\ u, & X > u. \end{cases}$$

In this case, if the loss is $X > u$, the insurance company pays an amount u. In other words the insurer pays a maximum amount of u on a claim. The following lemma gives the entropy expression of the random variable $V(u)$.

Lemma 5 *The entropy of the random variable $V(u)$ is given by*

$$H_{V(u)} = H_X + \int_u^\infty f_X(x) \ln f_X(x)\,dx - \bar{F}_X(u) \ln \bar{F}_X(u),$$

where H_X is the entropy of X, $\bar{F}_X(x)$ is the survival function of X and u is the liability limit.

4 The entropy of the multiplicative ruin problem

Consider random claim occurences modelled as homogeneous Poisson process $N[t]$, $t > 0$ with intensity $\lambda(t) > 0$, for all t, independent of the claim sizes Z_k, $k = 1, 2$ from a distribution $f(x)$. Suppose also that the effect of each claim is sequentially multiplicative. Then the total claim amount up to time t is

$$X(t) = \prod_{k=1}^{N[t]} Z_k.$$

The ruin problem is $Y(u,t) = P(\max_{s>0, s<t} X(s) > u + cs)$, where u is the initial capital and c is the constant premium.

Lemma 6 *The entropy of the total claim amount $X(t)$ equals*

$$H(X(t)) = \lambda(t) H_Z.$$

5 Conclusions

The role of this paper is double. We first investigate the use of entropy as an uncertainty measure for loss distributions and in general for actuarial mathematics. Secondly, we investigate the effect on the entropy computation of deductibles and policy (liability) limits. These situations frequently appear in actuarial practice. Thus the general and the specific for each loss model analytic expressions are given in relation to the entropy of the original data.

A characterization property of the per payment entropy is that is not depend on d if and only if the hazard function $\lambda_X(x)$ is constant. This is the case of exponential distribution and per payment insurance is the only case of partial insurance coverage that this property holds. Several other interesting properties of the per payment entropy were presented. However, the effect of deductible and policy limits on entropy is not easily interpretable while the expressions for the entropy computations are very complicated and differ among the models. In some cases it is impossible to find an analytic expression.

Entropy, as a measure of uncertainty and information, is useful for the study and evaluation of actuarial models. A well known method for determining uknown probability models is the method of *maximum entropy*. Due to this method, starting with some moments, which provide the only information at hand for the model, one can select the appropriate model by maximizing its entropy. This method is widely used in several sciences such as economics, accounting, biology, medicine, ecology, actuarial science etc. [11]. In actuarial science, the maximum entropy principle can be used in order to find the probability distribution of the number of claims on an insurance company in a time interval, to find the distribution of catastrophic events etc. Thus the entropy of the cases studied above are intresting and can be used in a maximum entropy context. Thus analytic expressions of entropy for the cases studied in the present paper, would be of interest.

References

[1] Abramowitz, M. and Stegun, I.A. (1972). *Handbook of Mathematical Functions*, National Bureau of Standards, Applied Mathematics Series, No. 55.

[2] Ahmed, N.A. and Gokhale, D.V. (1989). Entropy expressions and their estimators for multivariate distributions, *IEEE Transactions on Information Theory*, IT-35, 688-692.

[3] Arndt, C. (2001). *Information Measures - Information and its Description in Science and Engineering*, Springer - Verlang, Berlin.

[4] Belzunce, F., Navarro, J., Ruiz, J. and Aguila, Y. (2004). Some results on residual entropy function, *Metrika*, 59, 147-161.

[5] Darbellay, G.A. and Vajda, I. (2000). Entropy expressions for multivariate continuous distributions, *IEEE Transactions on Information Theory*, IT-46, 709-712.

[6] Di Crescenzo, A. and Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions, *Journal of Applied Probability*, 39, 434-440.

[7] Ebrahimi, N. (1996). How to measure uncertainty in the residual life distributions, *Sankhya A*, 58, 48-57.

[8] Ebrahimi, N., Kirmani, S.N.U.A., and Soofi, E.S. (2007). Multivariate dynamic information, *Journal of Multivariate Analysis*, 98, 328-349.

[9] Guerrero-Cusumano, J. (1996). An asymptotic test of independence for multivariate t and Cauchy random variables with applications, *Information Sciences*, 92, 33-45.

[10] Harris, B. (1982). Entropy, in: S. Kotz, N.L. Johnson (Eds.), *Encyclopedia of Statistical Sciences*, vol. 2, Wiley, New York, 512-516.

[11] Kapur, J.N. (1989). *Maximum-Entropy Models in Science and Engineering*, Wiley, New York.

[12] Klugman, S., Panjer, H.H., and Willmot, G.E. (2004). *Loss Models: From Data to Decisions*, New York: John Wiley and Sons, Inc, 2nd Edition.

[13] Nadarajah, S. and Zografos, K. (2003). Formulas for Rényi information and related measures for univariate distributions, *Information Sciences*, 155, 119-138.

[14] Nadarajah, S. and Zografos, K. (2005). Expressions for Rényi and Shannon entropies for bivariate distributions, *Information Sciences*, 170, 173-189.

[15] Pardo, L. (2006). *Statistical Inference Based on Divergence Measures*, Chapman & Hall/CRC.

[16] Sachlas, A. and Papaioannou, T. (2008). Computation of entropy under inflation and truncation from above, Proceedings of the 21st Annual Conference of Greek Statistical Institute, 275-282 (in Greek).

[17] Shannon, C.E. (1948). A mathematical theory of communication, *Bell System Technical*, 27, 379-423.

[18] Zografos, K. (2008). On some entropy and divergence type measures of variability and dependence for mixed continuous and discrete variables, *Journal of Statistical Planning and Inference*, 138, 3899-3914.

[19] Zografos, K. and Nadarajah, S. (2005). Expressions for Rényi and Shannon entropies for multivariate distributions, *Statistics & Probability Letters*, 71, 71-84.