# ESTIMATING PROBABLE MAXIMUM LOSS WITH ORDER STATISTICS

MARGARET E. WILKINSON

*Abstract*

In the past there has been much discussion about the definition of probable maximum loss (PML), but little attention has been given to its quantification. This paper introduces the concept of order statistics as a tool to use in estimating the PML. Two different approaches, that of $X_{(n)}$, the largest sample value, and that of quantiles, lead to six specific methods to estimate the PML. Three of the methods require sample data, two of the methods require assumptions about the underlying distribution of the population and the frequency, and one of the methods requires only estimates of the mean and variance of the population and of the frequency. All six methods are illustrated using a particular size of loss distribution. The methods work equally as well if the distribution of size of loss as a percentage of value is available.

## INTRODUCTION

The term PML is usually used in connection with property insurance, but it can also be applied to liability insurance. In fact, there is some controversy over whether the appropriate term, from a risk management viewpoint, is probable maximum loss, maximum possible loss, estimated maximum loss or one of many other similar phrases.

McGuinness [1] offers two definitions:

> "The probable maximum loss *for a property* is that proportion of total value of the property which will equal or exceed, in a stated proportion of all cases, the amount of loss from a specified peril or group of perils.
>
> "The probable maximum loss *under a given insurance contract* is that proportion of the limit of liability which will equal or exceed, in a stated proportion of all cases, the amount of any loss covered by a contract."

The first definition is pertinent to insureds and risk managers, while the second is pertinent to underwriters. These definitions were later combined by McGuinness [2] into one generalized definition:

> "The PML for a specified financial interest is that proportion of the total value of the interest which will equal or exceed, in a stated proportion of all cases, the amount of any financial loss to the interest from a specified event or group of events."

A guest reviewer [3] of McGuinness's paper, who is an underwriter, offered the following observations:

> "It is true that the definitions may vary between underwriters when put down in words, but I feel strongly that there is a universal meaning as to the end result which all underwriters expect PML to accomplish. . . . PML, no matter how you define it, is simply *Probable Maximum Loss*. It is neither *foreseeable* nor *possible* loss—rather, it is the maximum loss which *probably* will happen when, and if, the peril insured against actually occurs."

The concept of probable maximum loss used in this paper will not be defined separately from the definitions implied by the various measures to be discussed.

The PML depends upon (i) estimates of the likelihood that losses of various sizes will occur, (ii) the amount of losses and associated probabilities that the insured is willing to accept, and (iii) the amount of losses and associated probabilities that the underwriter is not willing to accept. Thus, the insured and the underwriter can have different estimates of the PML for the same loss exposure.

## ORDER STATISTICS

Let $X_1, X_2, \ldots, X_n$ denote a random sample from a population with continuous cumulative distribution function $F_X$. Since $F_X$ is continuous, the probability of any two sample values being equal is zero. Consequently, there exists a unique ordered arrangement of the sample. Let $X_{(1)}$ denote the smallest member of the set, $X_{(2)}$ the second smallest, etc. Then

$$X_{(1)} < X_{(2)} < \cdots < X_{(n)}$$

and these are called the *order statistics* from the random sample $X_1, X_2, \ldots, X_n$. For $1 \leq r \leq n$, $X_{(r)}$ is called the $r^{th}$ *order statistic.*

Order statistics are particularly useful for studying certain phenomena because quite a few of the results concerning the properties of $X_{(r)}$ and the properties of functions of some subset of the order statistics are distribution-free. If an inference is distribution-free, assumptions regarding the underlying population are not necessary. The distribution-free inference is based on a random variable which has a distribution independent of the underlying population's distribution.

## GENERAL RESULTS CONCERNING $X_{(n)}$

$X_{(n)}$ is the largest value of the sample. This is a good place to start since probable maximum loss is the worst loss likely to happen.

### *Distribution of* $X_{(n)}$

The cumulative distribution function of $X_{(n)}$ is given by

$$\begin{aligned} F_{X_{(n)}}(x) &= \Pr\{X_{(n)} \leq x\} \\ &= \Pr\{\text{all } X_i \leq x\} \\ &= F_X{}^n(x) \qquad (1) \end{aligned}$$

since the $X_i$'s are independent. The corresponding density function is found by differentiating (1). It is easily verified that

$$f_{X_{(n)}}(x) = nf_X(x)F_X{}^{n-1}(x) \qquad (2)$$

where $f_X$ is the density function corresponding to $F_X$.

### *Moments of* $X_{(n)}$

The exact moments of $X_{(n)}$ can be derived from the following equation:

$$E(X_{(n)}{}^k) = \int_{-\infty}^{\infty} x^k f_{X_{(n)}}(x)dx$$

$$= \int_{-\infty}^{\infty} nx^k f_X(x)F_X{}^{n-1}(x)dx. \tag{3}$$

This requires a specified distribution $F_X$ and is of limited practical value due to the complexity of the integral involved.

There are large-sample approximations for the mean and variance of $X_{(n)}$ that are easily calculable. The approximations require two facts.

1. If $U_{(r)}$ denotes the $r^{th}$ order statistic from a uniform distribution over the interval (0,1), then

   $X_{(r)} = F_X^{-1}(U_{(r)})$.

2. The Taylor's series expansion of a function $g(z)$ about a point $\mu$ is

   $$g(z) = g(\mu) + \sum_{i=1}^{\infty} \frac{(z - \mu)^i}{i!} g^{(i)}(\mu)$$

   where $g^{(i)}(\mu) = \frac{d^i g(z)}{dz^i} \Big|_{z=\mu}$ .

   This series converges if

   $$\lim_{n\to\infty} \frac{(z - \mu)^n}{n!} g^{(n)}(z_1) = 0$$

   for $\mu < z_1 < z$.

The first requirement is due to the probability integral transformation and is proved in various statistical texts [4]. The second requirement is the standard Taylor's series expansion.

If the Taylor's series expansion is rewritten for a random variable $Z$ with mean $\mu$, and the expected value of both sides is taken, the result is

$$E[g(Z)] = g(\mu) + \frac{\text{var }(Z)}{2!} g^{(2)}(\mu)$$

$$+ \sum_{i=3}^{\infty} \frac{E[(Z - \mu)^i]}{i!} g^{(i)}(\mu).$$

So, a first approximation to $E[g(Z)]$ is $g(\mu)$, and a second approximation is

$$g(\mu) + \frac{\text{var}\,(Z)}{2!}\, g^{(2)}(\mu).$$

To find similar approximations for $\text{var}[g(Z)]$, form the difference $g(Z) - E[g(Z)]$, square it and take the expected value. The result is

$$\text{var}[g(Z)] = \text{var}(Z)\, [g^{(1)}(\mu)]^2 - \frac{1}{4}\, [g^{(2)}(\mu)]^2 \text{var}^2(Z) + E[h(Z)]$$

where $E[h(Z)]$ involves third or higher central moments of $Z$ [5]. A first approximation to $\text{var}[g(Z)]$ is $\text{var}(Z)[g^{(1)}(\mu)]^2$, and a second approximation is $\text{var}(Z)[g^{(1)}(\mu)]^2 - (1/4)\, [g^{(2)}(\mu)]^2\, \text{var}^2(Z)$.

In order to apply these results to $X_{(n)}$, $g$ is defined so that

$$g(u_{(n)}) = x_{(n)} = F_X^{-1}(u_{(n)})$$

where $u_{(n)} = F_X(x_{(n)})$. The appropriate moments [6] are

$$\mu = E[u_{(n)}] = n/(n+1)$$

and

$$\text{var}[u_{(n)}] = \frac{n}{(n+1)^2(n+2)}\,.$$

The derivatives needed [7] are

$$g^{(1)}(\mu) = \{f_X[F_X^{-1}\,(n/(n+1))]\}^{-1}$$

and

$$g^{(2)}(\mu) = -f_X'[F_X^{-1}\,(n/(n+1))]\{f_X[F_X^{-1}\,(n/(n+1))]\}^{-3}.$$

Substituting yields as first approximations:

$$E(X_{(n)}) \simeq F_X^{-1}\,(n/(n+1)) \tag{4}$$

$$\text{var}(X_{(n)}) \simeq \frac{n}{(n+1)^2(n+2)}\, \{f_X[F_X^{-1}\,(n/(n+1))]\}^{-2}. \tag{5}$$

Second approximations are similarly found by the appropriate substitutions.

*Distribution-Free Bounds for* E(X$_{(n)}$) [8]

If a variate $X$ has a finite variance, the expected value of $X_{(n)}$ can not be arbitrarily large even if the range of $X$ is unbounded.

From Equation (3), the expected value of $X_{(n)}$ is

$$E(X_{(n)}) = \int_{-\infty}^{\infty} nxF_X{}^{n-1}(x)f_X(x)\,dx.$$

Let $u = F_X(x)$ and standardize $X$ to have mean 0 and variance 1. This means

$$E(X_{(n)}) = \int_0^1 nx(u)u^{n-1}du,$$

$$\int_0^1 x(u)du = 0,$$

$$\int_0^1 [x(u)]^2du = 1,$$

where $x(u)$ indicates that $x$ is expressed as a function of $u$.

Schwartz's inequality states that

$$\int fg\ du \leq (\int f^2du \int g^2du)^{1/2}.$$

Let $f = x$ and $g = nu^{n-1} - 1$. Then

$$\int_0^1 x(nu^{n-1} - 1)du \leq \left(\int_0^1 x^2du \int_0^1 (nu^{n-1} - 1)^2du\right)^{1/2}.$$

Expanding yields

$$\int_0^1 xnu^{n-1}du - \int_0^1 x\ du$$

$$\leq \left(\int_0^1 x^2du\right)^{1/2} \left(\int_0^1 (n^2u^{2n-2} - 2nu^{n-1} + 1)du\right)^{1/2}.$$

Substituting for the various pieces gives

$$E(X_{(n)}) \leq \left(\int_0^1 (n^2u^{2n-2} - 2nu^{n-1} + 1)du\right)^{1/2}.$$

Hence

$$E(X_{(n)}) \leq (n - 1)/(2n - 1)^{1/2}.$$

If the mean and variance of the population are $\mu$ and $\sigma^2$, respectively, the result becomes

$$E(X_{(n)}) \leq \mu + (n - 1)\, \sigma/(2n - 1)^{1/2} \tag{6}$$

This result is distribution-free and requires only the knowledge of the mean and variance of the population, not its specific distribution.

## GENERAL RESULTS FOR QUANTILES

Probable maximum loss has been defined as the worst loss likely to happen. If the sample under consideration has an unreasonably large loss, then using $X_{(n)}$ to estimate the PML would be unreasonable. In this case, quantiles could be used. The quantile approach would also be preferred if the insured was willing to accept more risk or the underwriter wanted to accept less risk. "More risk" and "less risk" used in this context are comparable to the expected retained losses implied by using $X_{(n)}$ to estimate the PML.

A quantile of a continuous distribution $f_X(x)$ of a random variable $X$ is a real number which divides the area under the probability density function into two parts of specified amounts. Denote the $p^{th}$ quantile by $\kappa_p$ for $0 \leq p \leq 1$. Then $\kappa_p$ is defined as any real number solution to the equation

$$F_X(\kappa_p) = p.$$

It is assumed that there is a unique solution to this equation, as there would be if $F_X$ is strictly increasing.

### *Point Estimate for* $\kappa_p$ [9]

It can be shown that the $r^{th}$ order statistic is a consistent estimator of the $p^{th}$ quantile where $r/n = p$ remains fixed. A definition which provides a unique $X_{(r)}$ to estimate the $p^{th}$ quantile is to choose $r$ so that

$$r = \begin{cases} np & \text{if } np \text{ is an integer} \\ [np + 1] & \text{if } np \text{ is not an integer} \end{cases} \tag{7}$$

where $[x]$ denotes the greatest integer not exceeding $x$.

### *Distribution-Free Confidence Interval for* $\kappa_p$ [10]

Since consistency is only a large-sample property, it is desirable to have an interval estimate for $\kappa_p$ with a known confidence coefficient for a given sample size. The objective is to find two numbers $r$ and $s$, $r < s$, such that

$$P(X_{(r)} < \kappa_p < X_{(s)}) = 1 - \alpha$$

for some chosen number $0 < \alpha < 1$.

For all $r < s$,

$$P(X_{(r)} < \kappa_p < X_{(s)}) = P(X_{(r)} < \kappa_p) - P(X_{(s)} < \kappa_p).$$

Since $F_X$ is a strictly increasing function,

$$X_{(r)} < \kappa_p \text{ if and only if } F_X(X_{(r)}) < F_X(\kappa_p) = p.$$

Thus,

$$\begin{aligned} P(X_{(r)} < \kappa_p < X_{(s)}) &= P[F_X(X_{(r)}) < p] - P[F_X(X_{(s)}) < p] \\ &= \int_0^p n \binom{n-1}{r-1} x^{r-1}(1-x)^{n-r} dx \\ &\quad - \int_0^p n \binom{n-1}{s-1} x^{s-1}(1-x)^{n-s} dx. \end{aligned}$$

If this formula is integrated by parts the necessary number of times, the result is

$$P(X_{(r)} < \kappa_p < X_{(s)}) = \sum_{i=r}^{s-1} \binom{n}{i} p^i(1-p)^{n-i}. \tag{8}$$

This does not produce a unique solution for $r$ and $s$. The narrowest interval is produced when $X_{(s)} - X_{(r)}$ is minimized. Alternatively, $s - r$ could be minimized. Also, a confidence interval produced by

$$\sum_{i=r}^{s-1} \binom{n}{i} p^i(1-p)^{n-i} = 1 - \alpha$$

is distribution-free.

The formula derived above can also be argued directly. For any $p$, $X_{(r)} < \kappa_p$ if and only if at least $r$ of the sample values $X_1, X_2, \ldots, X_n$ are less than $\kappa_p$. The sample values are independent and can be classified according to whether they are less than $\kappa_p$. Thus, the $n$ random variables can be considered the result of $n$ independent trials of a Bernoulli variable with parameter $p$. The number of observations less than $\kappa_p$ then has a binomial distribution with parameter $p$.

## APPLICATION OF ORDER STATISTICS TO THE PML PROBLEM

The application of order statistics has various requirements depending on the approach taken. The PML can simply be estimated by $X_{(n)}$ if a reliable data set applicable to the particular problem is available. If the concern is to estimate the PML by using the expected value of $X_{(n)}$ or by constructing an interval around $X_{(n)}$ using the variance of $X_{(n)}$ and choosing the PML as the upper limit of this interval, the distribution of $X$, $F_X$, must be known (actually $F_X^{-1}$, $f_X$ and $f_X'$ are needed). If estimates of the mean and variance of $F_X$ are available, derived either theoretically or from a data set, then the upper bound for $E(X_{(n)})$ could be used as the PML. If a data set is available but, for various reasons, the quantile approach is preferred, only the order statistics themselves are necessary to produce either a point estimate for the quantile or a confidence interval for the quantile. In the former case, the PML would be the quantile; in the latter case, the PML would be the upper bound of the confidence interval.

The data set or theoretical distribution used in estimating PML must be adjusted for trend. As there are several excellent papers [11] available on various methods of adjusting for trend, this paper will assume such adjustment has been made.

### $X_{(n)}$ *as an Estimate for PML*

Exhibit I contains a list of 100 claims that are representative of a particular problem in which a PML estimate is needed. $X_{(n)}$ in this case is $X_{(100)}$ or \$576,525. Consequently the PML is \$576,525.

### $E(X_{(n)})$ *as an Estimate for the PML*

The use of $E(X_{(n)})$ as an estimate for the PML requires $F_X^{-1}$. Suppose it is assumed that the data has a lognormal distribution. The mean is \$212,521 and the standard deviation is \$110,506. The corresponding normal distribution has a mean of 12.14714 and a standard deviation of .48920. From Equation (4), the approximation for the expected value of $X_{(n)}$ is

$$E(X_{(n)}) \simeq \Lambda_X^{-1}(n/(n+1)) = e^{[\sigma Z^{-1}(n/(n+1)) + \mu]}$$

where $\Lambda_X$ is the lognormal distribution,
$Z$ is the standard normal distribution,
$\mu$ is the mean of the normal distribution, and
$\sigma$ is the standard deviation of the normal distribution.

If $n = 100$, the value of $Z^{-1}(.9901)$ is found from standard normal tables to be 2.33. The PML estimate is \$589,468.

### *The Upper Bound of an Interval Around* $\mathrm{E}(\mathrm{X}_{(\mathrm{n})})$ *Using* $var(\mathrm{X}_{(\mathrm{n})})$ *as an Estimate for the PML*

It is possible to choose $k$ so that

$$E(X_{(n)}) + k(\mathrm{var}(X_{(n)}))^{1/2}$$

produces a reasonable estimate of the risk that is acceptable. If the prior example is continued, the $\mathrm{var}(X_{(n)})$ can be approximated using Equation (5):

$$\mathrm{var}(X_{(n)}) \simeq [100/(101)^2(102)]\ (\lambda_X(589{,}468))^{-2}$$

where $\lambda_X$ is the density function corresponding to $\Lambda_X$. The formula for $\lambda_X$ is

$$\lambda_X(x) = \frac{1}{x\sigma(2\pi)^{1/2}}\, e^{\{-(1/2\sigma 2)(\ln x - \mu)2\}}.$$

The $(\mathrm{var}(X_{(n)}))^{1/2}$ is \$106,976 for this example. If $k$ is chosen to be 2.0, the PML estimate is \$803,420.

### *The Distribution-Free Upper Bound of* $\mathrm{E}(\mathrm{X}_{(\mathrm{n})})$ *as an Estimate for the PML*

The data shown in Exhibit I have a sample mean of \$212,521 and a sample standard deviation of \$110,506. Consequently,

$$E(X_{(100)}) \leq 212{,}521 + 99\ (110{,}506)/(199)^{1/2}.$$

The PML is thus \$988,044.

If sample data are not available, a mean, variance and number of claims could be chosen on some theoretical grounds and the upper bound calculated as shown above.

### $\kappa_\mathrm{p}$ *as an Estimate for the PML*

Suppose it is decided that the .95 quantile will be used as the PML. If the sample data from Exhibit I are used, $r$ is 95 (because $.95 \times 100 = 95$) and the PML $(X_{(95)})$ is \$434,449.

### *The Distribution-Free Upper Bound of* $\kappa_\mathrm{p}$ *as an Estimate for the PML*

The estimate of $\kappa_p$ for $p = .95$ based on the sample data is \$434,449. Now a confidence interval is desired around this estimate so that $\alpha = .10$. In other words, $r < s$ must be found so that

$$P(X_{(r)} < \kappa_p < X_{(s)}) = \sum_{i=r}^{s-1} \binom{n}{i} p^i(1 - p)^{n-i} = .90.$$

We should also minimize $s - r$. Exhibit II shows $X_{(i)}$ and

$$\binom{n}{i} p^i(1 - p)^{n-i} \text{ for } i = 90, 91, \ldots, 100.$$

There are two possibilities for $r$ and $s$:

$$P(X_{(91)} < \kappa_{.95} < X_{(99)}) = .934732$$

and

$$P(X_{(92)} < \kappa_{.95} < X_{(99)}) = .899831.$$

The second is closer to .90 and $s - r$ is 7. The first has an $s - r$ of 8. Even though the probabilities are so close, and the second probability is slightly less than .90, the second answer would be chosen because $s - r$ is minimized. The PML in this case is $X_{(99)}$ or \$563,899.

In the above six examples a particular size of loss distribution was assumed. The PML estimates for the sample data are summarized in Exhibit III. While these estimates vary considerably, this is due to differing data and loss aversion considerations. The methods presented work equally well if the distribution of size of loss as a percentage of value is available. The former is more correct for liability insurance or for property insurance if the population has the same property value as the insured. The latter is more correct for property insurance where the property values differ among properties.

## SUMMARY

This paper has presented two different approaches to the PML problem using order statistics: $X_{(n)}$ and quantiles. These approaches lead to six different methods for estimating the PML:

1. $X_{(n)}$,
2. $E(X_{(n)})$,
3. $E(X_{(n)}) + k(\text{var}(X_{(n)}))^{1/2}$,
4. distribution-free upper bound of $E(X_{(n)})$,
5. $X_{(r)}$ as an estimate of $\kappa_p$, and
6. distribution-free upper bound of $\kappa_p$.

Methods 1, 5 and 6 require sample data. Methods 2 and 3 require assumptions about $n$ and the underlying distribution of the population. Method 4 requires only estimates of $n$ and the mean and variance of the population. The choice of method would depend on availability of data, willingness to make assumptions about the underlying population, and the amount of losses and associated probabilities the insured is willing to accept or the underwriter is not willing to accept.

## REFERENCES

[1] McGuinness, John S., "Is 'Probable Maximum Loss' (PML) A Useful Concept?" *Proceedings of the Casualty Actuarial Society,* Vol. LVI, 1969, p. 31.

[2] McGuinness, John S., "Author's Review of Discussions in Volume LVI, Pages 40–48," *Proceedings of the Casualty Actuarial Society,* Vol. LVII, 1970, p. 107.

[3] Black, Edward B., "Discussion by Edward B. Black," *Proceedings of the Casualty Actuarial Society,* Vol. LVI, 1969, p. 46.

[4] In particular, see Gibbons, Jean D., *Nonparametric Statistical Inference,* New York, 1971, p. 23.

[5] Ibid., p. 35.

[6] Ibid., pp. 32–33.

[7] Ibid., p. 37.

[8] David, Herbert A., *Order Statistics,* New York, 1981, pp. 56–59.

[9] Gibbons, op. cit., pp. 40–41.

[10] Ibid., pp. 41–43.

[11] For example, see Rosenberg, S. and Halpert, A. "Adjusting Size of Loss Distributions for Trend," *Inflation Implications for Property-Casualty Insurance,* Casualty Actuarial Society, 1981, p. 458.

# EXHIBIT I

## ORDERED SAMPLE DATA

| $i$ | $X_{(i)}$ | $i$ | $X_{(i)}$ |
|---|---|---|---|
| 1 | $ 19,874 | 51 | $207,196 |
| 2 | 30,610 | 52 | 208,959 |
| 3 | 32,159 | 53 | 209,568 |
| 4 | 34,115 | 54 | 213,084 |
| 5 | 40,660 | 55 | 214,307 |
| 6 | 53,453 | 56 | 214,546 |
| 7 | 56,598 | 57 | 215,978 |
| 8 | 61,651 | 58 | 216,369 |
| 9 | 63,411 | 59 | 220,808 |
| 10 | 66,007 | 60 | 222,804 |
| 11 | 73,062 | 61 | 224,417 |
| 12 | 76,962 | 62 | 224,475 |
| 13 | 87,348 | 63 | 235,209 |
| 14 | 96,498 | 64 | 238,249 |
| 15 | 98,408 | 65 | 238,679 |
| 16 | 109,837 | 66 | 238,842 |
| 17 | 122,838 | 67 | 240,455 |
| 18 | 128,372 | 68 | 244,699 |
| 19 | 128,426 | 69 | 247,465 |
| 20 | 130,048 | 70 | 251,374 |
| 21 | 130,610 | 71 | 257,426 |
| 22 | 131,326 | 72 | 258,513 |
| 23 | 131,474 | 73 | 265,051 |
| 24 | 137,655 | 74 | 269,816 |
| 25 | 139,681 | 75 | 271,647 |
| 26 | 140,949 | 76 | 274,154 |
| 27 | 147,987 | 77 | 275,727 |
| 28 | 150,776 | 78 | 277,211 |
| 29 | 151,044 | 79 | 277,734 |
| 30 | 151,967 | 80 | 279,494 |

## EXHIBIT I

### Ordered Sample Data

| $i$ | $X_{(i)}$ | $i$ | $X_{(i)}$ |
|---|---|---|---|
| 31 | 152,219 | 81 | 280,721 |
| 32 | 153,388 | 82 | 293,728 |
| 33 | 154,619 | 83 | 302,641 |
| 34 | 157,065 | 84 | 308,771 |
| 35 | 162,956 | 85 | 311,612 |
| 36 | 169,142 | 86 | 314,410 |
| 37 | 170,262 | 87 | 319,722 |
| 38 | 171,988 | 88 | 323,711 |
| 39 | 173,391 | 89 | 327,927 |
| 40 | 174,049 | 90 | 331,179 |
| 41 | 175,689 | 91 | 345,130 |
| 42 | 180,406 | 92 | 368,095 |
| 43 | 182,223 | 93 | 371,194 |
| 44 | 183,399 | 94 | 396,911 |
| 45 | 190,532 | 95 | 434,449 |
| 46 | 195,658 | 96 | 440,639 |
| 47 | 197,482 | 97 | 447,171 |
| 48 | 199,788 | 98 | 482,259 |
| 49 | 203,310 | 99 | 563,899 |
| 50 | 205,796 | 100 | 576,525 |

mean = $212,521
standard deviation = $110,506

## EXHIBIT II

### BINOMIAL PROBABILITIES FOR $n = 100, p = .95$

| $i$ | $X_{(i)}$ | $\binom{100}{i}(.95)^i(.05)^{100-i}$ |
|---|---|---|
| 90 | 331,179 | .016716 |
| 91 | 345,130 | .034901 |
| 92 | 368,095 | .064871 |
| 93 | 371,194 | .106026 |
| 94 | 396,911 | .150015 |
| 95 | 434,449 | .180018 |
| 96 | 440,639 | .178143 |
| 97 | 447,171 | .139576 |
| 98 | 482,259 | .081182 |
| 99 | 563,899 | .031161 |
| 100 | 576,525 | .005921 |

## EXHIBIT III

### SUMMARY OF EXAMPLE PML CALCULATIONS

| Method | PML Estimate |
|---|---|
| 1. $X_{(n)}$ | $576,525 |
| 2. $E(X_{(n)})$ | 589,468 |
| 3. $E(X_{(n)}) + k(\text{var}(X_{(n)}))^{1/2}$ | 803,420 |
| 4.* upper bound of $E(X_{(n)})$ | 988,044 |
| 5. $X_{(r)}$ as an estimate of $\kappa_p$ | 434,449 |
| 6.* upper bound of $\kappa_p$ | 563,899 |

*These are distribution-free.