Introduction

The following material is no longer on the syllabus:

- Lesson 6, The lognormal distribution
- Lesson 17, Ruin theory
- Lesson 30, Estimating parameters of a lognormal distribution
- Lessons 33–36, Cox model and generalized linear model
- Section 37.2, exercises 37.11–37.15, Normal plots
- Section 61.3, exercises 61.23–61.28, Simulating stocks and options
- Lessons 63–65, Special methods for simulating stocks and options
- Sections 68.1 and 68.3, exercises 68.1–68.6, 68.9–68.13 Distortion risk measures and semi-variance
- Anything in lessons 66–67 relating to risk measures for discrete distributions. For example, Example 66D, exercises 66.1–66.2, Section 67.2, exercises 67.3–67.4. In addition, the risk measure called CTE in the previous syllabus is now called Tail VaR, or TVaR.

Also, the material on Approximations for Large Data Sets (Lesson 25) no longer uses α, β, and P_j notation. The only two cases discussed in the syllabus are what is called $\alpha = 1$, $\beta = 0$ and $\alpha = 0.5$, $\beta = 0.5$ in the old syllabus. The double-decrement formula—the displayed line in Section 25.2—is no longer mentioned in the syllabus.

The following practice exam questions are no longer on the syllabus:

1: 17, 23, 26, 32, 34, 40
2: 6, 19, 23, 29, 37
3: 14, 21, 28, 35, 36, 40
4: 12, 20, 22, 30, 35
5: 3, 10, 16, 24, 30, 40
6: 6, 22, 23, 24, 29, 34, 38
7: 1, 14, 24, 25, 26, 35

Officially, the only new material added to the syllabus is Fitting $(a, b, 1)$ distributions. However, the syllabus also moved from the second edition to the third edition of *Loss Models*, and includes a new section on Non-normal confidence intervals for maximum likelihood estimators. Therefore, both of these topics are provided below. Cross-references are to the Exam C/4 9th edition manual, but you shouldn't have too much difficulty adjusting them to your edition.

Non-Normal Confidence Intervals

Section 15.4 of *Loss Models*, which is new to the third edition, discusses non-normal confidence intervals for parameters.

The use of normal confidence intervals assumes the maximum likelihood estimator is normally distributed, which is true asymptotically but not for small samples, and that building separate confidence intervals for each parameter separately is optimal. An alternative method for building confidence intervals is to solve an inequality for the loglikelihood equation. The confidence interval consists of the k-dimensional region in which the loglikelihood function is greater than c for some constant c, where k is the number of parameters. This region is not necessarily cubical, unlike the normal confidence region.

As we'll learn later when we're studying likelihood ratio tests in lesson 35, if we want p confidence, c is selected to be the maximum loglikelihood minus $0.5w$, where w is the pth percentile of the chi-square distribution with k degrees of freedom. Once again, k is the number of parameters being estimated.

To illustrate non-normal confidence intervals, let's repeat the previous example.

EXAMPLE 0A You are given:

(i) A random variable has probability density function

$$f(x) = ax^{a-1} \qquad 0 \le x \le 1, a > 0$$

(ii) The parameter a is estimated by maximum likelihood.

(iii) A random sample of observations of X is 0.3, 0.6, 0.6, 0.8, 0.9.

Construct a 95% non-normal confidence interval for a.

ANSWER: There is only one parameter. We select a region in which $l(a) = 5\ln a + (a-1)\ln\prod_{i=1}^{5} x_i > c$ for some c. Since $\ln\prod_{i=1}^{5} x_i = -2.55413$, this reduces to the set

$$\{a \mid 5\ln a - 2.55413(a-1) > c\}$$

In our example, the maximum value of the loglikelihood function is

$$l(1.95762) = 5\ln 1.95762 + (0.95762)(-2.55413) = 0.91276$$

The 95th percentile of chi-square with one degree of freedom is 3.841, so we want $l(x) \ge 0.912768 - 0.5(3.841) = -1.0077$. The equation

$$l(x) = 5\ln x - 2.55413(x-1) = -1.0077$$

requires an iterative technique to solve. Using Excel's Solver, I obtained the interval $\boxed{\mathbf{(0.70206, 4.20726)}}$ for a. Compare this to the normal confidence interval $1.95762 \pm 1.96\sqrt{0.76645} = (0.24169, 3.67354)$. □

Note that in calculating c, the exact loglikelihood is needed. Multiplicative constants must not be dropped.

Since virtually any loglikelihood formula involves a combination of logs and polynomial terms, equating the loglikelihood to a constant usually requires a numerical technique. Therefore, there are very few possibilities for exam questions.

One possibility is to calculate a non-normal confidence interval for the parameter θ of a uniform distribution on $[0,\theta]$. Note that the asymptotic theory of this lesson does not apply to a uniform distribution, because its maximum likelihood estimator is not determined through differentiation. For a random variable having a uniform distribution on $[0,\theta]$, the MLE is the maximum observation. The variance of the maximum observation of a uniform distribution (if you want to construct a normal confidence interval) can be determined directly; see formula 17.2 on page 266. The likelihood function of a sample of n is $1/\theta^n$, with loglikelihood $-n\ln\theta$. You can set this equal to its maximum value minus c, and solve this for θ. Note that the likelihood is 0 for $\theta < \max x_i$, so any confidence interval for θ is a one-sided interval with left boundary $\max x_i$.

EXAMPLE 0B A random variable follows a uniform distribution on $[0,\theta]$. A sample of 100 observations of the random variable has maximum 50. The parameter θ is estimated using maximum likelihood.

Construct a 95% non-normal confidence interval for θ.

ANSWER: The loglikelihood function is $l(\theta) = -100\ln\theta$ for $\theta \ge 50$, and is maximized at $\theta = 50$ as $-100\ln 50 = -391.202$. Subtracting half of the 95th percentile of chi-square with one degree of freedom, or 0.5(3.841), from the maximum loglikelihood, we have

$$\begin{aligned} -100\ln x &\ge -391.202 - 0.5(3.841) = -393.123 \\ \ln x &\le 3.93123 \\ x &\le e^{3.93123} = 50.97 \end{aligned}$$

The confidence interval is $\boxed{\mathbf{[50, 50.97)}}$. □

Another possibility for an exam question is the confidence interval for the estimator of μ of a normal or lognormal distribution with fixed σ. However, as the next example shows, this is not too interesting.

EXAMPLE 0C A random variable X has a normal distribution with variance 100. A random sample of 25 observations has sample mean $\bar{x}$. The parameter μ is estimated from these observations using maximum likelihood.

Construct a 95% non-normal confidence interval for μ.

ANSWER: The likelihood and loglikelihood (expressed generally in terms of $\sigma = 10$ and $n = 25$) are

$$L(\mu) = \frac{e^{-\sum(x_i-\mu)^2/2\sigma^2}}{(\sigma\sqrt{2\pi})^n}$$
$$l(\mu) = -n\ln(\sigma\sqrt{2\pi}) - \frac{\sum(x_i-\mu)^2}{2\sigma^2}$$

The maximum likelihood estimate of μ is $\bar{x}$, which you can prove by differentiating $l(\mu)$ and setting the derivative equal to 0. So the maximum loglikelihood is

$$l(\bar{x}) = -n\ln(\sigma\sqrt{2\pi}) - \frac{\sum(x_i-\bar{x})^2}{2\sigma^2}$$

Now, suppose μ' is in the 95% confidence interval. In that case, we need $2|l(\mu') - l(\bar{x})| \le 3.841$. Note that in the expression $|l(\mu') - l(\bar{x})|$, the first summand $-n\ln(\sigma\sqrt{2\pi})$ cancels. We are left with

$$2|l(\mu') - l(\bar{x})| = \left|\frac{\sum\left((x_i-\mu')^2 - (x_i-\bar{x})^2\right)}{\sigma^2}\right|$$

Now,

$$\begin{aligned}\sum(x_i-\mu')^2 &= \sum(x_i-\bar{x}+\bar{x}-\mu')^2\\ &= \sum(x_i-\bar{x})^2 + 2\sum(x_i-\bar{x})(\bar{x}-\mu') + n(\bar{x}-\mu')^2\\ &= \sum(x_i-\bar{x})^2 + n(\bar{x}-\mu')^2\end{aligned}$$

because $\sum(x_i-\bar{x}) = n\bar{x} - n\bar{x} = 0$. So

$$2|l(\mu') - l(\bar{x})| = \frac{n(\bar{x}-\mu')^2}{\sigma^2}$$

In our case, $n = 25$ and $\sigma^2 = 100$. So we need

$$0.25(\bar{x}-\mu')^2 \le 3.841$$
$$|\bar{x}-\mu'| \le \sqrt{15.364} = 3.920$$

and that defines the non-normal confidence interval. Now note that the variance of the sample mean is $100/25 = 4$ whose square root is 2, so that the normal confidence interval would also be $\bar{x} \pm 2(1.96) = \bar{x} \pm 3.92$. So both confidence intervals are identical. But what did you expect for a normal random variable? □

A remaining possibility for an exam question would be to ask whether a specific point is in the non-normal confidence interval.

Fitting $(a,b,1)$ class distributions

We will use the same notation as in Section 10.2. Namely, p_k^M or p_k^T is the probability that the $(a,b,1)$ random variable is k in the modified or truncated distribution respectively, whereas p_k is the probability that the corresponding $(a,b,0)$ random variable, the one with the same a and b, is k.

Maximum likelihood has the following properties when fitting $(a,b,1)$ distributions:

1. The fitted probability of 0 will match the proportion of 0 in the sample. In other words

$$\hat{p}_0^M = \frac{n_0}{n}$$

   where n_0 is the number of observed zeros and n is the sample size.

   If the sample has any observations of 0, maximum likelihood will not fit a truncated distribution, since in such a distribution the likelihood of 0 is zero. The converse is true as well since the fitted probability of 0 matches the observed proportion.

2. The fitted mean will equal the sample mean. As a result of this and the first property, the conditional mean given that the variable is greater than 0 will equal the sample mean of the non-zero observations.

Numerical methods are needed to calculate the parameters other than p_0^M in almost every case. The textbook derives formulas, and none of these formulas allow solutions without a computer.

For a zero-modified Poisson, the formula is

$$\bar{x} = \frac{1-\hat{p}_0^M}{1-p_0}\lambda$$

and since $p_0 = e^{-\lambda}$, this is a mixed exponential/linear equation. Note that the right-hand side is the theoretical mean of the distribution.

For a zero-modified binomial, it is necessary to construct likelihood profiles for each m, and once again the sample mean equals the theoretical mean:

$$\bar{x} = \frac{1-\hat{p}_0^M}{1-p_0}mq$$

Since $p_0 = (1-q)^m$, this is a high degree polynomial. You can maximize the likelihood for $m \le 3$, but that's about it. The possibility $m = 1$ is only available if no observations are greater than 1, and is the same as fitting a Bernoulli which you can easily calculate maximum likelihood for, so that only leaves $m = 2$ or $m = 3$.

EXAMPLE 0D You have the following observations of a discrete random variable:

Value	Number of Observations
0	10
1	6
2	4

You are to fit these to a zero-modified binomial distribution using maximum likelihood.

Calculate the fitted value of q when $m = 2$ and when $m = 3$, and determine the resulting loglikelihoods.

ANSWER: As discussed above, $\hat{p}_0^M = n_0/n = 10/20 = 0.5$. Then for $m = 2$,

$$\begin{aligned}
\frac{1-0.5}{1-(1-q)^2}(2q) = \bar{x} &= \frac{6+4(2)}{20} = 0.7 \\
q &= 0.7(2q-q^2) = 1.4q - 0.7q^2 \\
0.7q &= 0.4 \\
\hat{q} &= \frac{4}{7}
\end{aligned}$$

The fitted probabilities are

$$p_0^M = 0.5$$
$$p_1^M = \frac{0.5\big(2q(1-q)\big)}{1-(1-q)^2}$$
$$p_2^M = \frac{0.5q^2}{1-(1-q)^2}$$

so the likelihood and loglikelihood are

$$L(q \mid m=2) = \frac{0.5^{20}\big(2q(1-q)\big)^6\big(q^2\big)^4}{\big(1-(1-q)^2\big)^{10}}$$
$$l(q \mid m=2) = 14\ln 0.5 + 14\ln q + 6\ln(1-q) - 10\ln\big(1-(1-q)^2\big)$$
$$= -9.70406 - 7.83462 - 5.08379 + 2.02941 = -20.5931$$

For $m=3$,

$$\frac{0.5}{1-(1-q)^3}(3q) = 0.7$$
$$1.5q = 0.7(3q - 3q^2 + q^3)$$
$$1.5 = 2.1 - 2.1q + 0.7q^2$$
$$0.7q^2 - 2.1q + 0.6 = 0$$
$$\hat{q} = \frac{2.1 - \sqrt{2.73}}{1.4} = 0.31981$$

The other solution to the quadratic is rejected since then $q > 1$.

The likelihood and loglikelihood are

$$L(q \mid m=3) = \frac{0.5^{20}\big(3q(1-q)^2\big)^6\big(3q^2(1-q)\big)^4}{\big(1-(1-q)^3\big)^{10}}$$
$$l(q \mid m=3) = 20\ln 0.5 + 10\ln 3 + 14\ln q + 16\ln(1-q) - 10\ln\big(1-(1-q)^3\big)$$
$$= -13.86294 + 10.98612 - 25.66462 - 2.78781 + 8.98789 = -22.3421$$

Since the maximum likelihood is less for $m=3$ than for $m=2$, it will continue to decline for higher values of m. The maximum likelihood fit is $\hat{m}=2$, $\hat{q}=4/7$. □

For a zero-modified negative binomial, two parameters in addition to p_0^M must be fitted. We needed numerical methods even for a non-modified negative binomial, and certainly need them here. Even when r is known numerical methods are usually required to solve the rational equation equating the mean to the sample mean, with the exception of special values of r. One particularly easy case is $r=1$, which characterizes a zero-modified geometric.

Example 0E You have the following observations of a discrete random variable:

Value	Number of Observations
0	10
1	6
2	4

You are to fit these to a zero-modified geometric distribution using maximum likelihood.

Determine $\hat{\beta}$.

ANSWER: Maximum likelihood reduces to method of moments. The probability of zero is the sample probability, or 0.5. Then the mean is matched to the sample mean of 0.7:

$$\begin{aligned}(1-p_0^M)(1+\beta)&=0.7\\0.5(1+\beta)&=0.7\\\hat{\beta}&=\boxed{\mathbf{0.4}}\end{aligned}$$ □

Because of the difficulty in fitting $(a,b,1)$ distributions, as well as the rarity of any $(a,b,1)$ distribution questions on past exams, I think the likelihood of questions from this topic, which was added to the syllabus starting with the Fall 2009 exam, is low.

Exercises

1. Losses follow a Pareto distribution with parameters $\theta = 1000$ and α. Five observed losses were

   500 500 800 2000 10,000

   The parameter α is estimated using maximum likelihood.

   Select the smallest of the following numbers which is contained in a 95% non-normal confidence interval for α.

   (A) 0.2 (B) 0.3 (C) 0.4 (D) 0.5 (E) 0.6

2. X follows a lognormal distribution with parameters μ and $\sigma = 2$. A sample of 25 observations has geometric mean e^5. The parameter μ is estimated using maximum likelihood.

   Construct 95% asymptotic normal and 95% non-normal confidence intervals for μ.

3. In a sample of ten policyholders who submitted at least one claim in a year, nine of them submitted exactly one claim and the tenth one submitted n claims.

   The data were fitted to a zero-truncated Poisson distribution using maximum likelihood. The estimated value of λ is 0.54986.

   Determine n.

4. You are given the following accident data from 100 insurance policies:

Number of accidents	Number of policies
0	70
1	16
2	8
3	6
4+	0

   The data are fitted to a zero-modified negative binomial distribution with $r = -0.5$.

   Determine the fitted value of β.

Solutions

1. By the Pareto MLE shortcut, the estimate $\hat{\alpha}$ is

$$K = 5\ln 1000 - \ln(1500^2)(1800)(3000)(11{,}000) = 34.53878 - 39.43400 = -4.89522$$
$$\hat{\alpha} = \frac{-5}{-4.89522} = 1.02140$$

and the loglikelihood function is

$$l(\alpha) = 5\ln\alpha + 5\alpha\ln 1000 - (\alpha+1)\sum\ln(1000 + x_i) = 5\ln\alpha - 4.89522\alpha - 39.434$$

So for α' to be in the interval, we need

$$|5\ln(\alpha/\alpha') - 4.89522(\alpha - \alpha')| \le 1.92$$

Plugging in 0.4:

$$5\ln(1.02140/0.4) - 4.89522(1.02140 - 0.4) = 1.645$$

so 0.4 is in the interval. Plugging in 0.3,

$$5\ln(1.02140/0.3) - 4.89522(1.02140 - 0.3) = 2.594$$

so 0.3 is out of the interval. The answer is **0.4**.

2. The loglikelihood function is

$$l(\mu) = -0.5n\ln 2\pi\sigma^2 - \sum\ln x_i - \frac{\sum(\ln x_i - \mu)^2}{2\sigma^2} \qquad (*)$$

Differentiating once and setting equal to 0, we find

$$\frac{\sum(\ln x_i - \mu)}{\sigma^2} = 0$$

so $\hat{\mu} = (\sum\ln x_i)/n$. In our case,

$$\hat{\mu} = \frac{\sum\ln x_i}{n} = \ln\sqrt[n]{\prod x_i} = \ln e^5 = 5$$

Differentiating again, we get n/σ^2 as the information matrix, so the asymptotic variance is $\sigma^2/n = 4/25 = 0.16$. The asymptotic normal confidence interval is $5 \pm 1.96\sqrt{0.16} =$ **(4.216, 5.784)**.

Let μ' be a point in the 95% non-normal confidence interval. Twice the difference between the loglikelihood of μ' and of $\hat{\mu}$, since the first two summands of (*) are the same for both, is

$$\frac{\sum\left((\ln x_i - 5)^2 - (\ln x_i - \mu')^2\right)}{4} = \frac{25(5-\mu')^2}{4}$$

where the simplification of the numerator is discussed in the solution to Example 0C. So we need

$$\frac{25(5-\mu)^2}{4} \le 3.841$$
$$5 - \mu \le 0.4\sqrt{3.841} = 1.96(0.4)$$

and we end up with the same confidence interval as the asymptotic normal confidence interval.

3. The maximum likelihood estimate of the expected value is the sample mean. The expected value of a zero-truncated Poisson is $\lambda/(1-e-\lambda)$, and $0.54986/(1-e^{-0.54986}) = 1.3$. It follows that the sample mean is 1.3, or $n = \boxed{\mathbf{4}}$.

4. Maximum likelihood sets $\hat{p}_0^M$ equal to the proportion of observed 0's, or 0.7. Then we match the fitted mean to the sample mean, which is $\big(16(1)+8(2)+6(3)\big)/100 = 0.5$.

$$
\begin{aligned}
0.3\left(\frac{-0.5\beta}{1-(1+\beta)^{0.5}}\right) &= 0.5 \\
-0.3\beta &= 1-(1+\beta)^{0.5} \\
0.3\beta + 1 &= (1+\beta)^{0.5} \\
0.09\beta^2 + 0.6\beta + 1 &= 1+\beta \\
0.09\beta - 0.4 &= 0 \\
\hat{\beta} &= \boxed{\mathbf{\frac{40}{9}}}
\end{aligned}
$$