STAT3035/8035
Tutorial 8

Marco Li

Contact: qingyue.li@anu.edu.au

Outline

Approximating Compound Distributions

- Recall from last week:
  - The *CDF* of S, $G(s) = \sum_{n=0}^{\infty} F^{*(n)}(s) p_N(n)$ is very complicated, and the calculation of n-fold convolution $F^{*(n)}(s)$ is difficult
  - We learnt the *mgf* of S: $m_S(t) = m_N[\ln\{m_X(t)\}]$. However, some distributions for X_i have no *mgf* (e.g., Pareto or Log-Normal)
  - $S = \sum_{i=1}^{N} X_i$ is a (random) sum of independent random variables, maybe CLT applies?
- Normal approximation to $G(s)$
  - Form of approximation:

$$G(s) = \Pr(S \leq s) \approx \Phi\left\{\frac{s - E(S)}{\sqrt{\mathrm{Var}(S)}}\right\}$$

  - Advantage: Only need $E(S) = \nu\mu_1$ and $\mathrm{Var}(S) = \nu\mu_2 + \mu_1^2\left(\tau^2 - \nu\right)$ (Don't need $p_N(n)$ or $f_X(x;\theta)$'s, just first two moments)
  - Accuracy of Normal Approximation depends on skewness
    - Standardised coefficient of skewness $\rho_S = \frac{\mathrm{Skew}(S)}{\{\mathrm{Var}(S)\}^{3/2}}$
    - Rule of Thumb: If $\rho_S < 0.5$, Normal Approximation is good

Approximating Compound Distributions

- Translated Gamma approximation to $G(s)$
  - Idea: Gamma distribution can introduce skewness to our approximation
  - If $Y \sim G\left(\alpha_g, \theta_g\right)$, then $X = Y + k \sim transG\left(k, \alpha_g, \theta_g\right)$
  - If $X = Y + k \sim transG\left(k, \alpha_g, \theta_g\right)$, then

$$E(X) = k + \alpha_g \theta_g; \quad Var(X) = \alpha_g \theta_g^2; \quad \rho_X = \frac{2}{\sqrt{\alpha_g}}$$

  - Approximate $G(s)$ by $F_X\left(s; k, \alpha_g, \theta_g\right)$ such that

$$E(S) = k + \alpha_g \theta_g; \quad Var(S) = \alpha_g \theta_g^2; \quad \rho_S = \frac{2}{\sqrt{\alpha_g}}$$

[Method of moments]

Outline

Question 1

(Compound Negative Binomial Distribution, Normal Approximation and Compound Poisson Approximation) Suppose that S represents the aggregate claim amount from a portfolio which makes a negative binomial number of claims with parameters $k = 20$ and $q = 0.5$ and that these claims are Gamma distributed with shape parameter $\alpha = 3$ and scale parameter $\theta = 0.05$.
(a) Estimate the probability that S is less than 4.5 using the normal approximation.
(b) Repeat this exercise assuming that the portfolio makes a Poisson number of claims with rate parameter $\lambda = 20$. Compare your approximate answer to the "exact" answer 0.9654. Furthermore, compare your approximate answers here with those from part (a). Discuss your findings.

Compound Negative Binomial Distributions

$$
\begin{aligned}
E(S) &= k\mu_1(1-q)/q \\
\mathrm{Var}(S) &= k(1-q)\left\{q\mu_2 + (1-q)\mu_1^2\right\}/q^2 \\
\mathrm{Skew}(S) &= k(1-q)\left\{q^2\mu_3 + 3q(1-q)\mu_2\mu_1 + 2\left(1-2q+q^2\right)\mu_1^3\right\}/q^3
\end{aligned}
$$

Solution 1

Solution 1

(a) We have $\mathbb{E}X_i = \mu_1 = \alpha\theta = 3(0.05) = 0.15$ and $\mathbb{E}X_i^2 = \mu_2 = \alpha(\alpha+1)\theta^2 = 3(4)\left(0.05^2\right) = 0.03$. So, from Section 4.2.3 of the course notes, we have:

$$\begin{aligned}\mathbb{E}S &= k\left(q^{-1}-1\right)\mu_1 = 20(2-1)(0.15) = 3\\ \mathbb{V}S &= k\left(q^{-1}-1\right)\left\{\mu_2 + \left(q^{-1}-1\right)\mu_1^2\right\} = 20(2-1)\left\{0.03 + (2-1)\left(0.15^2\right)\right\} = 1.05\end{aligned}$$

Thus,

$$\mathbb{P}(S \leq 4.5) = \mathbb{P}\left(\frac{S-3}{\sqrt{1.05}} \leq \frac{4.5-3}{\sqrt{1.05}}\right) \approx \Phi\left(\frac{4.5-3}{\sqrt{1.05}}\right) = \Phi(1.46385) = 0.9283825$$

(b) In the case of a Poisson number of claims, we have:

$$\begin{aligned}\mathbb{E}S &= \lambda\mu_1 = 20(0.15) = 3\\ \mathbb{V}S &= \lambda\mu_2 = 20(0.03) = 0.6\end{aligned}$$

Thus,

$$\mathbb{P}(S \leq 4.5) = \mathbb{P}\left(\frac{S-3}{\sqrt{0.6}} \leq \frac{4.5-3}{\sqrt{0.6}}\right) \approx \Phi\left(\frac{4.5-3}{\sqrt{0.6}}\right) = \Phi(1.9365) = 0.9735968$$

As given in the question, the "exact" value is 0.9654, indicating that the normal approximation is rather accurate in this case. Moreover, we see that the approximate probability in this case is noticeably higher than in the case of negative binomial claim numbers, despite the fact that the expected number of claims is the same in each case. This is due to the greater variance in claim numbers under the negative binomial model, which indicates that larger total claim amounts are somewhat more likely in this model than under the Poisson claim number model.

Question 2

(Translated Gamma Approximation) Suppose that X has a Gamma distribution with parameters α and θ.
(a) Suppose that $k = 2\alpha$ is a positive integer. Show that $Y = 2X/\theta$ has a chi-squared distribution with k degrees of freedom. Note that the chi-squared distribution pdf is as follows:

$$f(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k/2)-1} e^{-x/2}.$$

Question 2

(Translated Gamma Approximation) Suppose that X has a Gamma distribution with parameters α and θ.
(a) Suppose that $k = 2\alpha$ is a positive integer. Show that $Y = 2X/\theta$ has a chi-squared distribution with k degrees of freedom. Note that the chi-squared distribution pdf is as follows:

$$f(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k/2)-1} e^{-x/2}.$$

Solution: We have $Y = 2\theta^{-1}X$, so that $X = \frac{1}{2}\theta Y$. Thus, a simple change of variable shows that:

$$\begin{aligned}
f_Y(y) &= f_X\left(\frac{1}{2}\theta y\right)\left|\frac{dx}{dy}\right| \\
&= \frac{1}{\theta^\alpha \Gamma(\alpha)}\left(\frac{1}{2}\theta y\right)^{\alpha-1} e^{-\theta^{-1}\left(\frac{1}{2}\theta y\right)}\left|\frac{1}{2}\theta\right| \\
&= \frac{1}{\theta^\alpha \Gamma(\alpha)}\left(\frac{1}{2}\right)^{\alpha} \theta^\alpha y^{\alpha-1} e^{-\frac{1}{2}y} \\
&= \frac{1}{2^\alpha \Gamma(\alpha)} y^{\alpha-1} e^{-\frac{1}{2}y} \\
&= \frac{1}{2^{k/2}\Gamma(k/2)} y^{(k/2)-1} e^{-\frac{1}{2}y}
\end{aligned}$$

where $k = 2\alpha$. Thus, Y has a chi-squared distribution with $k = 2\alpha$ degrees of freedom.

Question 2

(b) Suppose that you wanted to approximate the probability that a compound distributed quantity S exceeded some value using the translated Gamma method, but you only had chi-squared tables at your disposal. Discuss how your result from part (a) would be useful.

Question 2

(b) Suppose that you wanted to approximate the probability that a compound distributed quantity S exceeded some value using the translated Gamma method, but you only had chi-squared tables at your disposal. Discuss how your result from part (a) would be useful.

Solution: As noted in lectures, if $Y \sim G(\alpha_g, \theta_g)$ and $X = Y + k$ is the translated Gamma approximation to S, we have:

$$
\begin{aligned}
\Pr(S \le s) &\approx \Pr(X \le s) \\
&= \Pr(Y \le s - k) \\
&= \Pr\left\{2\theta_g^{-1} Y \le 2\theta_g^{-1}(s-k)\right\} \\
&= \Pr\left\{\chi^2_{2\alpha} \le 2\theta_g^{-1}(s-k)\right\}
\end{aligned}
$$

Question 2

(c) Suppose that $\mathbb{E}S = 10, \mathbb{V}S = 30$ and $\text{Skew}(S) = 270$. Use the translated Gamma approximation and chi-squared tables to estimate $\mathbb{P}(S \leq 20.9)$. Also, estimate the value of s which solves the equation $\mathbb{P}(S \leq s) = 0.99$.

Chi-square Distribution Table

d.f.	.995	.99	.975	.95	.9	.1	.05	.025	.01
1	0.00	0.00	0.00	0.00	0.02	2.71	3.84	5.02	6.63
2	0.01	0.02	0.05	0.10	0.21	4.61	5.99	7.38	9.21
3	0.07	0.11	0.22	0.35	0.58	6.25	7.81	9.35	11.34
4	0.21	0.30	0.48	0.71	1.06	7.78	9.49	11.14	13.28
5	0.41	0.55	0.83	1.15	1.61	9.24	11.07	12.83	15.09
6	0.68	0.87	1.24	1.64	2.20	10.64	12.59	14.45	16.81
7	0.99	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21

Solution 2

Solution 2

(c) From the given information, we can calculate the coefficient of skewness for S as $\rho_S = \frac{270}{30^{3/2}} = 1.643$. Thus, the appropriate translated Gamma distribution for use in approximating S has parameters k, α_g and θ_g which solve the equations:

$$1.643 = \rho_S = \frac{2}{\sqrt{\alpha_g}}, \quad 30 = \alpha_g \theta_g^2, \quad 10 = k + \alpha_g \theta_g$$

Solving the first equation shows that $\alpha_g = 4/(1.643)^2 = 1.482$, which implies that $\theta_g = \sqrt{30/1.482} = 4.5$ and $k = 10 - 4.5(1.482) = 3.333$. So, letting $Y \sim G(1.482, 4.5)$, we have

$$\begin{aligned} \mathbb{P}(S \leq 20.9) \approx \mathbb{P}(Y + 3.333 \leq 20.9) = \mathbb{P}(Y \leq 17.567) = \mathbb{P}\left\{2\left(4.5^{-1}\right) Y \leq 7.808\right\} \\ \approx \mathbb{P}\left(\chi_3^2 \leq 7.808\right) \approx 0.95 \end{aligned}$$

[NB : The actual value of $\mathbb{P}(Y \leq 17.567)$ is 0.9512.] Similarly, the 99 th percentile of S is the value $s_{0.99}$ which satisfies:

$$0.99 = \mathbb{P}\left(S \leq s_{0.99}\right) \approx \mathbb{P}\left(Y \leq s_{0.99} - 3.333\right) \approx \mathbb{P}\left\{\chi_3^2 \leq 2\left(4.5^{-1}\right)\left(s_{0.99} - 3.333\right)\right\}$$

Now, the 99th percentile of a chi-squared distribution with three degrees of freedom is 11.34, so that we can estimate the 99th percentile of S as the solution to:

$$2\left(4.5^{-1}\right)\left(s_{0.99} - 3.333\right) = 11.34 \quad \Longrightarrow \quad s_{0.99} = 28.85$$

Question 3

Suppose that S represents the aggregate claim amount from a portfolio which makes a Poisson number of claims with rate parameter $\lambda = 20$ and that these claims are Weibull distributed with parameters $\theta = 100$ and $\gamma = 0.5$

(a) Use the normal approximation to estimate the probability that S exceeds 750,000.

(b) Use the normal approximation to estimate the 98th percentile of the distribution of S.

Moments of Weibull Distribution

$$\mathbb{E}X_i^k = \theta^{k/\gamma}\Gamma\left(\frac{k}{\gamma} + 1\right)$$

Question 3

Suppose that S represents the aggregate claim amount from a portfolio which makes a Poisson number of claims with rate parameter $\lambda = 20$ and that these claims are Weibull distributed with parameters $\theta = 100$ and $\gamma = 0.5$
(a) Use the normal approximation to estimate the probability that S exceeds 750,000.
(b) Use the normal approximation to estimate the 98th percentile of the distribution of S.

Moments of Weibull Distribution

$$\mathbb{E}X_i^k = \theta^{k/\gamma}\Gamma\left(\frac{k}{\gamma} + 1\right)$$

Solution: (a) We have $\mu_1 = \mathbb{E}X_i = \theta^2\Gamma(3) = \left(100^2\right)(2!) = 20000$ and $\mu_2 = \mathbb{E}X_i^2 = \theta^4\Gamma(5) = \left(100^4\right)(4!) = 2400000000$. Therefore, $\mathbb{E}S = \lambda\mu_1 = 400000$ and $\mathbb{V}S = \lambda\mu_2 = 48000000000$. So, using the normal approximation, we have:

$$\mathbb{P}(S > 750000) \approx 1 - \Phi\left(\frac{750000 - 400000}{\sqrt{48000000000}}\right) = 1 - 0.9449 = 0.0551$$

(b) We know that the 98th percentile of the standard normal distribution is $z_{0.98} = 2.05375$ Therefore, the approximate 98th percentile of S in this case is:

$$s_{0.98} \approx \mathbb{E}S + z_{0.98}\sqrt{\mathbb{V}S} = 400000 + 2.05375\sqrt{48000000000} = 849954$$

Question 3

(c) Repeat parts (a) and (b) using the translated Gamma approximation method.

$$\mathbb{P}\left(\chi^2_4 > 8.933\right) = 0.0628 \qquad \chi^2_4(0.98) = 11.668$$

$$\mathbb{P}\left(\chi^2_5 > 8.933\right) = 0.1118 \qquad \chi^2_5(0.98) = 13.388$$

Solution 3

Solution 3

(3) In order to use the translated Gamma approximation, we will need $\mu_3 = \mathbb{E}X_i^3 = \theta^6\Gamma(7) = \left(100^6\right)(6!) = 7.2 \times 10^{14}$. This implies that the coefficient of skewness of S is

$$\rho_S = \frac{\text{Skew}(S)}{\mathbb{V}S^{3/2}} = \frac{\lambda\mu_3}{(\lambda\mu_2)^{3/2}} = \frac{\mu_3}{\mu_2^{3/2}\sqrt{\lambda}} = \frac{7.2 \times 10^{14}}{\left(24 \times 10^8\right)^{3/2}\sqrt{20}} = \frac{7.2 \times 10^2}{24^{3/2}\sqrt{20}} = 1.369$$

Therefore, the appropriate parameters for the translated Gamma approximation are $\alpha_g = 4/\left(1.369^2\right) = 2.133$, $\theta_g = \sqrt{48000000000/2.133} = 150000$ and $k = 400000$ - $2.133(150000) = 80000$. So

$$\mathbb{P}(S > 750000) \approx \Pr(Y + k > 750000) = \Pr(Y > 670000) = 0.0742$$

[NB : $\mathbb{P}\left\{\chi_4^2 > 2\left(150000^{-1}\right)(670000)\right\} = \mathbb{P}\left(\chi_4^2 > 8.933\right) = 0.0628$ and $\mathbb{P}\left(\chi_5^2 > 8.933\right) = 0.1118$, and since $2\alpha_g = 4.266$, we might approximate $\mathbb{P}(Y > 670000)$ by interpolation as $0.0628 + 0.266(0.1118 - 0.0628) = 0.0758$.] Similarly, the 98th percentile of $S, s_{0.98}$, satisfies:

$$0.98 = \mathbb{P}\left(S \le s_{0.98}\right) \approx \mathbb{P}\left(Y + k \le s_{0.98}\right) = \mathbb{P}\left(Y \le s_{0.98} - 80000\right)$$

Now, the 98th percentile of Y can be calculated as 910078.8 [NB : $\chi_4^2(0.98) = 11.668$ and $\chi_5^2(0.98) = 13.388$, so that we could approximate the 98th percentile of Y by interpolation as: $\frac{1}{2}150000\{11.668 + 0.266(13.388 - 11.668)\} = 909414$].
Therefore, $s_{0.98} \approx 910078.8 + 80000 = 990078.8$. Note that both the probability and the quantile are somewhat different than the estimates based on the normal approximation, and this is not surprising since the coefficient of skewness for S in this case was quite large, indicating that the distribution of S is quite skewed and thus not amenable to very accurate normal approximation.