Inventory Optimization Using Models for Count Data

Julian Sampedro

2020 (updated 2025)

Overview

(*i*) Investigation of daily sales of two products in a fictional online store.

(*ii*) Analysis of stock levels to meet the daily demand based on historical data (300 observations).

(*iii*) Comparision of different statistical models, namely:

- Poisson (P)
- Negative Binomial (NB)
- Zero-inflated Poisson (ZIP)
- Zero-inflated Negative Binomial (ZINP)

(*iv*) Discussion of results and suggest recommendations/actions to be implemented.

Notes: *For Poisson and Negative Binomial models, the parameter of interest will be estimated by maximum liklihood.*

Data and models

The Poisson (P) distribution is common for modeling count data. Let x_i denote the number of units sold on day i. Assuming that x_i obey a Poisson distribution with parameter $\lambda \in [0, \infty)$ (mean and dispersion), we have:

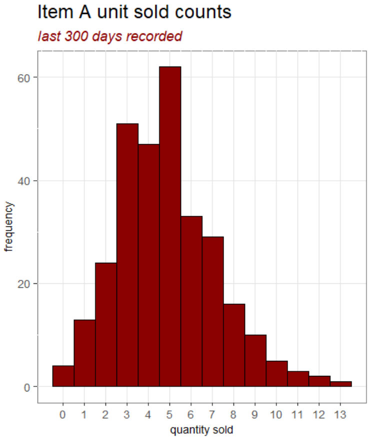
$$x_i \sim P(\lambda), \quad i = 1, \dots, n.$$

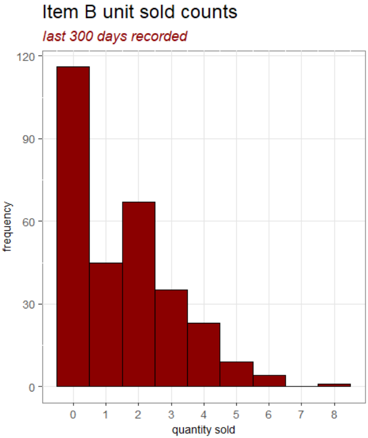
The Negative Binomial (NB) distribution is suitable for modelling count data that exhibit overdispersion. Now if we assume that x_i obey a Negative Binomial distribution with dispersion parameters $r > 0$ and probability of success $p \in [0, 1]$.

$$x_i \sim NB(r, p), \quad i = 1, \dots, n.$$

Notes: *Inventory optimization is common in retail analytics. For products with variable daily demand recorded as count data (our case) understocking can lead to lost sales, whereas overstocking increases holding costs. In our case, $i = 700$. The Bayesian modelling approach will be introduced later.*

Histograms





Note: *As shown on those histograms, the distribution of quantity sold for product A appears to be Poisson or Negative Binomial in case of overdispersion. For product B, there is an excess of* 0 *which suggest that a zero-inflated Poisson model might yiel better estimates.*

Poisson Maximum Likelihood Estimation

With daily sales records over $n = 300$ days and a total of sales of $\sum_{i=1}^{300} x_i = 1390$ (product A) and we obtain:

$$\hat{\lambda}_{\mathrm{MLE}} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{1451}{300} \approx 4.836667$$

Then our fitted Poisson model is $P(4.836667)$, where 4.836667 simply corresponds to the expected daily sales under the Poisson assumption.

A quick overdispersion check by computing $\phi = s_x^2/\bar{x}$ indicates that the distribution of quantity sold for product A is slightly overdispersed (1.650025). This suggest at least a comparison against a Negative Binomial model.

Negative Binomial Maximum Likelihood Estimation

For item A with daily sales records over 300 days and estimating r by $\hat{r}$ using the Method of Moments,

$$\hat{r}_{MOM} = \frac{\bar{x}^2}{s_x^2 - \bar{x}} \approx 41.4291, \text{with } \bar{x} = \sum_{i=1}^{300} x_i \text{ and } s_x^2 = \sqrt{\frac{\sum_{i=1}^{300}(x_i - \bar{x})^2}{300}},$$

our MLE for p is therefore

$$\hat{p}_{MLE} = \frac{300 \cdot 41.4291}{1451 + 300 \cdot 41.4291} \approx 0.895459.$$

Thus our second model (NB) is: $NB(41.4291, 0.8955)$ with mean:

$$\hat{\mu}_{mle} = \hat{r}_{MOM}\left(\frac{1 - \hat{p}_{MLE}}{\hat{p}_{MLE}}\right) = 4.8367$$

Zero-Inflated Models MLE

Zero-inflated models (here ZIP and ZINB) account for excess zeros in the data. They can be described as two components discrete mixture distribution, one being a distribution of interest (Poisson, ...) and a point mass at zero. For ZIP, the PMF is:

$$P(X = x \mid \pi, \lambda) = \begin{cases} \pi + (1-\pi)e^{-\lambda}, & x = 0, \\ (1-\pi)\frac{\lambda^x e^{-\lambda}}{x!}, & x = 1, 2, \ldots \end{cases}$$

where π is the zero-inflation probability and λ is the mean of the Poisson component. Now, for the ZINB PMF, we have:

$$P(X = x \mid \pi, n, p) = \begin{cases} \pi + (1-\pi)\, p^n, & x = 0, \\ (1-\pi) \binom{n+x-1}{x} p^n (1-p)^x, & x = 1, 2, \ldots \end{cases}$$

where π is the structural-zero probability (zero inflation), n is the dispersion (size) parameter and p is the success probability. In R, we use the package **pscl** to perform ML estimation as GLMs with intercept only.

Results

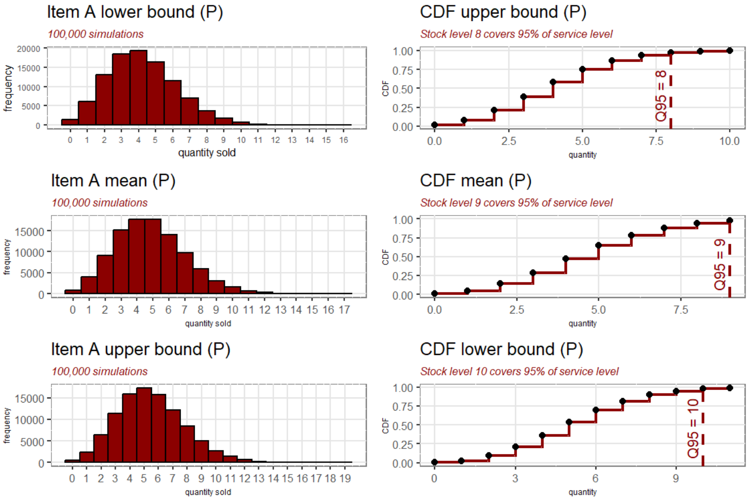
Table: Item A

	Estimate	*95% CI*
Poisson (P)	**4.84**	**4.23 – 5.55**
Negative Binomial (NB)	4.84	4.23 – 5.45
Zero-Inflated Poisson (ZIP)	4.86	4.25 – 5.48
Zero-Inflated NB (ZINB)	4.86	4.52 – 5.21

Table: Item B

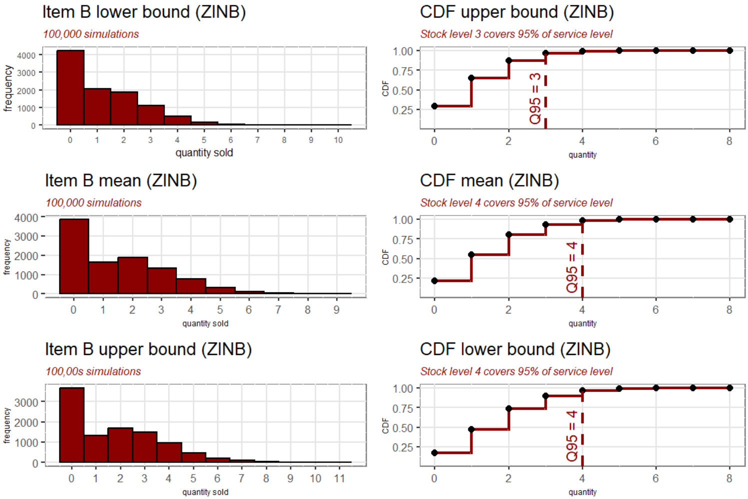
	Estimate	*95% CI*
Poisson (P)	1.51	1.70 – 1.85
Negative Binomial (NB)	1.51	1.70 – 1.85
Zero-Inflated Poisson (ZIP)	2.19	1.78 – 2.59
Zero-Inflated NB (ZINB)	**2.19**	**1.78 – 2.59**

CDF Estimation of Stock levels (A)



Note: *Consider lead-time*

CDF Estimation of Stock levels (B)



Note: *Consider lead-time*

Discussion of the results

(*i*) The demand for Item A is higher and stocking 9 will cover 95% of the service level (demand) on average or 8 for slow season.

(*ii*) The demand for Item B is lower (premium product) and stocking 4 will cover 95% of the service level (demand) on average.

(*iii*) Limitation: we did not include lead-time in our analysis.

(*iv*) What we used: maximum likelihood estimation, monte carlo simulations and cumulative distribution functions, visualisation, with the R programming language.

(*v*) Further possible work: consider lead-time (if known), joint analysis (for example using other statistical models), Bayesian version using hierarchical models, dynamic implementation in an interactive dashboard.

References

Roberto Rossi, Steven Prestwich, Armagan Tarim, Brahim Hnich, 2014. Confidence-based optimisation for the newsvendor problem under binomial, Poisson and exponential demand.

Rizzo, M.L. , 2019. Statistical Computing with R, Second Edition (2nd ed.). Chapman and Hall/CRC.

Course notes.

Appendix: Poisson ML Estimation

Starting from the Pobability Mass Function (PMF):

$$P(X = x \mid \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots,$$

for n daily sales observations $x_1, \ldots, x_n$, the likelihood is:

$$\mathcal{L}(\lambda \mid \mathbf{x}) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda} \prod_{i=1}^{n} \frac{1}{x_i!},$$

and thus the log-likelihood is:

$$\ell(\lambda \mid \mathbf{x}) = \sum_{i=1}^{n} x_i \ln(\lambda) - n\lambda - \sum_{i=1}^{n} \ln(x_i!).$$

Taking the derivative of the log-likelihood with respect to λ, setting it equal to 0, and solving for λ yields:

$$\frac{\partial \ell(\lambda \mid \mathbf{x})}{\partial \lambda} = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n = 0 \quad \Rightarrow \quad \hat{\lambda}_{\text{MLE}} = \frac{\sum_{i=1}^{n} x_i}{n} (= \bar{x})$$

Fisher information, asymptotic variance, and 95% asymptotic confidence interval:

$$I_n(\lambda) = \frac{n}{\lambda}, \quad \text{Var}(\hat{\lambda}_{\text{MLE}}) \approx \frac{\hat{\lambda}_{\text{MLE}}}{n},$$

$$\hat{\lambda}_{\text{MLE}} \pm z_{0.975} \sqrt{\frac{\hat{\lambda}_{\text{MLE}}}{n}}, \text{ with } z_{0.975} \approx 1.96$$

Appendix: Negative Binomial ML Estimation

The PMF is:

$$P(X = x \mid r, p) = \binom{r+x-1}{x} p^r (1-p)^x, \quad x = 0, 1, 2, \ldots,$$

where $\binom{r+x-1}{x} = \frac{\Gamma(r+x)}{\Gamma(r)\,\Gamma(x+1)}$ and $\Gamma()$ is the gamma function. For n daily observations $x_1, \ldots, x_n$, we have

$$\mathcal{L}(r, p \mid \mathbf{x}) = \prod_{i=1}^{n} \frac{\Gamma(r + x_i)}{\Gamma(r)\,\Gamma(x_i + 1)}\; p^{nr}\; (1-p)^{\sum_{i=1}^{n} x_i},$$

and for the log-likelihood, we have:

$$\ell(r, p \mid \mathbf{x}) = \sum_{i=1}^{n} \ln\left(\Gamma(r + x_i)\right) - n \ln(\Gamma(r)) - \sum_{i=1}^{n} \ln\left(\Gamma(x_i + 1)\right) + nr \ln(p) + \sum_{i=1}^{n} x_i \ln(1-p)$$

Taking the derivative of the log-likelihood with respect to p, setting it equal to 0 and solving for p yields:

$$\frac{\partial \ell(r, p \mid \mathbf{x})}{\partial p} = nr\frac{1}{p} - \sum_{i=1}^{n} x_i \frac{1}{1-p}$$

$$nr\frac{1}{p} - \sum_{i=1}^{n} x_i \frac{1}{1-p} = 0 \quad \Leftrightarrow \quad \hat{p}_{\mathbf{MLE}} = \frac{\mathbf{nr}}{\mathbf{nr} + \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{x_i}}$$

Fisher information for the sample, asymptotic variance and 95% asymptotic confidence interval:

$$I_n(p) = \frac{nr}{p^2(1-p)}, \qquad \mathrm{Var}(\hat{p}_{MLE}) \approx \frac{1}{I_n(\hat{p}_{MLE})},$$

$$\hat{p}_{MLE} \pm z_{0.975}\sqrt{\frac{1}{I_n(\hat{p}_{MLE})}}, \quad z_{0.975} \approx 1.96$$