3èmes Rencontres STAT à l'UBS

24-25 mars 2016

Van-Cuong DO and Evans Gouno
Université de Bretagne Sud, FRANCE

Introduction

Bonjour à toutes et à tous. Je suis très heureux d'être ici aujourd'hui pour vous présenter mon travail de thèse que j'effectue sous la direction de Monsieur Evans GOUNO.
It is my pleasure standing here today giving you the presentation of my work under the supervision of Mr. Evans GOUNO.
Relying of the work of **??**, we introduce the H-B distribution which appears as a natural conjugate prior for Bayesian analysis of the Power Law Process. Then we investigate some properties of H-B distribution in order to make elicitation of the hyper-parameters. Finally, we conduct Bayesian estimation of the parameters using different priors and make comparison between them and Maximum Likelihood Estimation.

What is PLP ?

The stochastic point process $\{N(t), t \geq 0\}$ is called PLP if it is a Non-homogeneous Poisson Process (NHPP) with intensity function

$$m(t) = \frac{\beta}{\alpha^\beta} t^{\beta-1}.$$

That means, the numbers of events in a time interval $(s, t]$, say $N(s, t)$ follows the Poison distribution with parameter

$$\lambda(s, t) = \int_s^t m(u)du = (t/\alpha)^\beta - (s/\alpha)^\beta.$$

So, on average, the number of events in the time interval $(0, t]$ is

$$M(t) = E(N(t)) = \int_0^t m(u)du = (t/\alpha)^\beta.$$

What is PLP ?

In order to make statistical inferences on $N(t)$, we need to estimate the two parameters α, β.
If $\beta > 1$ then the intensity $m(t)$ increases, the PLP is getting more events overtime. If $\beta < 1$ then the intensity $m(t)$ decreases, the PLP is getting less events overtime. In case of $\beta = 1$, the PLP degenerates to the HPP with the constant rate of events.

Duane Plot

In 1964, Duane found that when he plotted the data of failure dates collecting from repairable systems, they followed up straight lines. With the mathematical view, the Power Law Process fit well in explaining that phenomenon.

Life data from the electronics system

i= Failure number	t_i= Failure time	t_i/i=cum MTBF
1	0,1	0,10
2	5,6	2,80
3	18,6	6,20
4	19,5	4,88
5	24,2	4,84
6	26,7	4,45
7	45,1	6,44
8	45,6	5,70
9	75,7	8,41
10	79,7	7,97
11	98,6	8,96
12	120,1	10,01
13	161,8	12,45
14	180,6	12,90
15	190,8	12,72

Life data from the electronics system

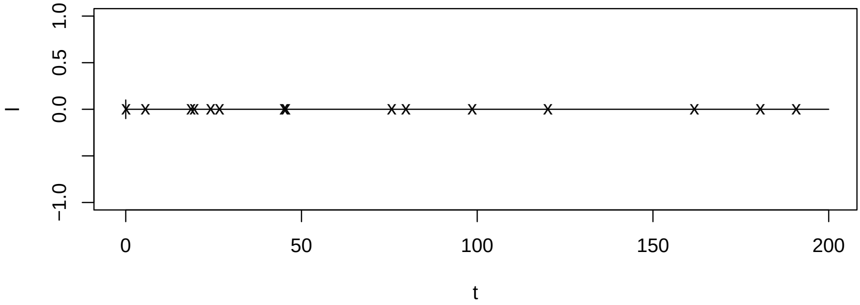
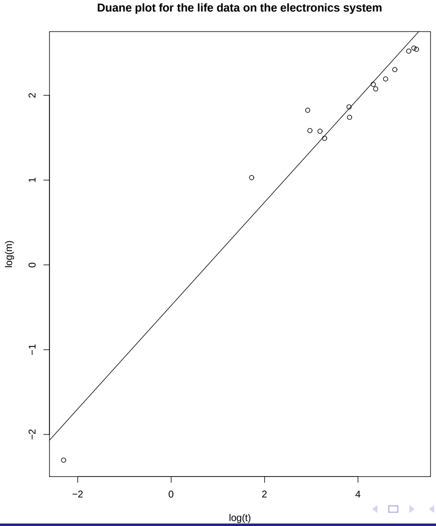


FIGURE: *Visualizing the life data from the electronics system*

Duane plot for graphical test



How to explain the Duane Plot

Suppose that the observation $(t_1, t_2, \ldots, t_n)$ follows the PLP. The MTBF is

$$K(t) = \frac{t}{M(t)} = \alpha^{\beta}.t^{1-\beta}$$
$$\Leftrightarrow \log(K(t)) = \log(\alpha^{\beta}) + (1-\beta)\log(t) \quad (1)$$

Therefore, $\{\log(t_i), \log(t_i/i); i = 1, \ldots, n\}$ fit the linear model with the slope $1-\beta$ and the intercept $\log(\alpha^{\beta})$

Describing the successive dates of events of the PLP

Let's consider the sequence of successive dates of events, say $(T_1, T_2, \ldots, T_n)$, of the PLP $\{N(t), t \geq 0\}$, where T_i is the random variable referring to the date of the i-th event.
It can be shown that T_1 has **Weibull distribution** with scale parameter α and shape parameter β because its survival function has form

$$S_{T_1}(t) = P(T_1 > t) = P(N(t) = 0)$$
$$= \exp\left\{-\int_{t_{i-1}}^{t} \frac{\beta}{\alpha}\left(\frac{s}{\alpha}\right)^{\beta-1} ds\right\} = \exp\left\{-\left(\frac{t}{\alpha}\right)^{\beta}\right\}. \qquad (2)$$

Describing the successive dates of events of the PLP

Now we can find the conditionally distribution of random variable T_i given the successive dates of events that happened before $(T_{i-1} = t_{i-1}, \ldots, T_1 = t_1)$. For $t \geq t_{i-1}$ we have

$$S(t \mid T_{i-1} = t_{i-1}, \ldots, T_1 = t_1) = P(T_i > t \mid T_{i-1} = t_{i-1}, \ldots, T_1 = t_1)$$
$$= P(N(t_{i-1}, t) = 0) = \exp\left\{ -\int_{t_{i-1}}^{t} \frac{\beta}{\alpha} \left(\frac{s}{\alpha}\right)^{\beta-1} ds \right\}$$
$$= \exp\left\{ -\left[\frac{t}{\alpha}\right]_{t_{i-1}}^{t} \right\} = \exp\left\{ -\frac{1}{\alpha^{\beta}}[t^{\beta} - t_{i-1}^{\beta}] \right\}. \quad (3)$$

So T_1 has **left-truncated Weibull distribution at the left point t_{i-1}.**

The likelihood function of PLP

Let's find the likelihood of the observation $(t_1, t_2, \ldots, t_n)$ (in this paper, we only talk about the failure truncation case).

$$f(t_1, t_2, \ldots, t_n \mid \alpha, \beta) = f_{T_1}(t_1).f_{T_2 \mid T_1 = t_1}(t_2). \ldots .f_{T_n \mid T_{i-1} = t_{i-1}, \ldots, T_1 = t_1}(t_n)$$

With the parameters (α, β) of intensity form $m(t) = \frac{\beta}{\alpha^\beta} t^{\beta - 1}$, the corresponding likelihood of the PLP is

$$L(\alpha, \beta) = \alpha^{-n\beta}.\beta^n \left(\prod_{i=1}^{n} t_i \right)^{\beta - 1} . \exp \left\{ - \left(\frac{t_n}{\alpha} \right)^\beta \right\} \qquad (4)$$

Maximum Likelihood Estimation

It is not difficult to find the Maximum Likelihood Estimators of both α, β. In fact, they have closed form without any numerical calculation.

$$\hat{\beta}_{MLE} = \frac{n}{\sum_{i=1}^{n} \log(y/T_i)}, \tag{5}$$

$$\hat{\alpha}_{MLE} = \frac{y}{n^{1/\hat{\beta}}}. \tag{6}$$

Because

$$E(\hat{\beta}_{MLE}) = \frac{n}{n-2}\beta$$

so $\hat{\beta}_{MLE}$ is biased estimator of β. It is useful to note that $2n\beta/\hat{\beta}_{MLE}$ has chi-square distribution with $2(n-1)$ degrees of freedom so we can conduct hypothesis testing for parameter β.

Draw-backs of frequentist approach and the Advantages of Bayesian approach

However, there is no pivotal quantity for the ML estimator of parameter α and numerical integration is needed. That is why we name α the nuisance parameter. In fact, this is one of the drawback of the frequentist approach. That is when the Bayesian approach come in !
Frequentist base on the fact that there are infinitive repetitive observations in the same condition. It could not be true in the real life. Each data we observe can be drawn by different parameters. Therefore, it is reasonable to consider parameters as random variables, not the given fixed values.

Draw-backs of frequentist approach and the Advantages of Bayesian approach

The Bayesian approach provides a natural, unified environment for carrying out the estimation and prediction process based on finite-sample calculations rather than the large-sample approximation often needed for the frequentist case. Moreover, it can readily incorporate any strong prior information in the inference process if one such is available. Most importantly, in the specific case of inference for the PLP, the Bayesian approach take advance to the frequentist framework. In the context of observing the failure process of a repairable system, two types of inspection schemes are typically adopted in practice.

Draw-backs of frequentist approach and the Advantages of Bayesian approach

The schemes, refered to as "time truncation" and "time truncation", closely resemble "Type I" and "Type II" censoring, respectively, in that the process terminates either at a predetermined time or a number of failures. The inference for the two sampling schemes (time-truncated and failure-truncated) are intrinsically different in the frequentist case. However, from the Bayesian perspective both failure truncation and time truncation data can be handled in the same manner and result in the same type of posterior inference on α and β in contrast to the classical frequentist approach in which each case must be treated separately and different types of results are obtained.

Draw-backs of frequentist approach and the Advantages of Bayesian approach

An important, relevant aspect is that in many situations collection of failure data on repairable systems can be too expensive or time consuming so that a limited number of failures can be observed. Hence, Bayes methods may be desirable as they allow for prior information to be incorporated into inferential procedure. (For references, see Bar-Lev et al. (1991), Crow (1982), Sen, A. (2002), Guida et al. (1989)).

Non-Informativ Prior for both α and β

For convenience in writing, we use these notation

$$\underline{t} = (t_1, \ldots, t_n), u = \prod_{i=1}^{n} t_i, v = \log(t_n^n/u). \tag{7}$$

According to the Jefreys' rule, a non-informative joint prior for α and β is

$$\pi(\alpha, \beta) \propto (\alpha\beta)^{-1}. \tag{8}$$

The likelihood is

$$L(\alpha, \beta) = \alpha^{-n\beta} \beta^n u^{\beta-1}. \exp\left\{-\left(\frac{y}{\alpha}\right)^{\beta}\right\}. \tag{9}$$

Non-Informativ Prior for both α and β

Applying Bayes' rule, the posterior is

$$\pi(\alpha, \beta \mid \underline{t}) = K(\underline{t}).\alpha^{-(n\beta+1)}\beta^{n-1}u^{\beta}.\exp\left\{-\left(\frac{t_n}{\alpha}\right)^{\beta}\right\}, \tag{10}$$

where

$$K(\underline{t}) = \frac{v^{n-1}}{\Gamma(n)\Gamma(n-1)}. \tag{11}$$

Non-Informativ Prior for both α and β

The marginal posterior density for β can be obtained by integrating out α

$$\pi(\beta \mid \underline{t}) = \frac{w^{n-1}}{\Gamma(n-1)} \beta^{n-2} e^{-v\beta}. \tag{12}$$

That means, the marginal posterior of β is Gamma distribution with parameters $(n-1, v)$. The Bayesian Estimation of parameter β is then

$$\hat{\beta}_{Bayes1} = E(\beta \mid \underline{t}) = \frac{n-1}{v} \tag{13}$$

Unfortunately, the Bayesian Estimation of parameter α does not has closed form and must be done numerically by MCMC.

Two independent Gamma distributions as conjugate prior

Olivera, M. D. Colosimo, E. A. and Gilardoni, G. L. (2011) proposed the new re-parametrization of the PLP in terms of (θ, β) where $\theta = M(t_n) = (t_n/\alpha)^\beta$. The likelihood becomes

$$L(\theta, \beta) \propto (\theta^n e^{-\theta}) \times (\beta^n e^{-v\beta}). \quad (14)$$

It means that θ and β are independent, so the natural conjugate family is the product of two Gamma distributions $\theta \sim Gamma(a, b)$, $\beta \sim Gamma(c, d)$

$$\pi(\theta, \beta) = \pi(\theta) \times \pi(\beta) \propto \theta^{a-1} e^{-b\theta} \times \beta^{c-1} e^{-d\beta}. \quad (15)$$

The posterior is then

$$\pi(\theta, \beta \mid \underline{t}) \propto \theta^{a+n-1} e^{-(b+1)\theta} \times \beta^{c+n-1} e^{-(d+v)\beta}. \quad (16)$$

Bayesian Estimators of α and β

That means $\theta \mid \underline{t} \sim Gamma(a+n, b+1)$, $\beta \mid \underline{t} \sim Gamma(c+n, d+v)$. Assuming a quadratic loss, the Bayesian estimators are the expectation of the posterior distributions. The Bayesian estimator of parameter β is then

$$\hat{\beta}_{Bayes2} = E(\beta \mid \underline{t}) = \frac{c+n}{d+v}, \tag{17}$$

$$\hat{\theta}_{Bayes2} = E(\theta \mid \underline{t}) = \frac{a+n}{b+1}. \tag{18}$$

Relation between Maximum Likelihood Estimators and Bayesian Estimators

Remark that $E(\beta) = c/d, \hat{\beta}_{MLE} = n/v$ and $E(\theta) = a/b, \hat{\theta}_{MLE} = n$. There is a relation between Maximum Likelihood Estimators and Bayesian Estimators :

$$\hat{\beta}_{Bayes2} = E(\beta \mid \underline{t}) = p\hat{\beta}_{MLE} + (1-p)E(\beta), \tag{19}$$

$$\hat{\theta}_{Bayes2} = E(\theta \mid \underline{t}) = q\hat{\theta}_{MLE} + (1-q)E(\theta) \tag{20}$$

where

$$p = \frac{v}{v+d}, q = \frac{1}{b+1} \tag{21}$$

Elicitation of the hyper-parameters

The hypothesis of independent Gamma distributions priors make it easy to elicitate the hyper-parameters. If we have the prior information about the expectation and variance of θ, say $E(\theta) = m_1, Var(\theta) = v_1$; the expectation and variance of β, say $E(\beta) = m_2, Var(\beta) = v_2$ then

$$a = \frac{m_1^2}{v_1}, b = \frac{m_1}{v_1}, \tag{22}$$

$$c = \frac{m_2^2}{v_2}, d = \frac{m_2}{v_2}. \tag{23}$$

Narural Conjugate Prior

Relying on the work of Huang and Bier (1998), we propose the re-parametrization of the PLP in terms of (λ, β) where $\lambda = 1/\alpha^\beta$, then the likelihood becomes

$$L(\lambda, \beta) \propto \lambda^n \beta^n u^\beta \exp\left\{-\lambda t_n^\beta\right\}. \tag{24}$$

The natural conjugate prior family should be in the form

$$\pi(\lambda, \beta) \propto \lambda^{a-1} \beta^{a-1} c^\beta \exp\left\{-b\lambda t_n^\beta\right\} \tag{25}$$

This conjugate prior looks like an unfamiliar distribution. We introduce here the new distribution named as H-B distribution.

The H-B distribution

Definition

A bivariate random variable $(X, Y) \in \mathbb{R}^+ \times \mathbb{R}^+$ is said to be distributed as the H-B distribution with four-parameter (a, b, c, d) if it has density of the form

$$f_{X,Y}(x, y) = K\ (xy)^{a-1} c^y \exp\{-bd^y x\} \tag{26}$$

where $a, b, c, d > 0$ and such that $c > d^a$; $K = [b \log(d^a/c)]^a / \Gamma(a)^2$.

We denote : $(X, Y) \sim$ H-B(a, b, c, d)

Contour plot and density plot of the H-B distribution

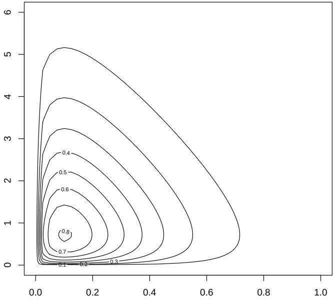


FIGURE: Contour plot of the H-B distribution with $a = 1.5$, $b = 5$, $c = 0.5$ and $d = 1$

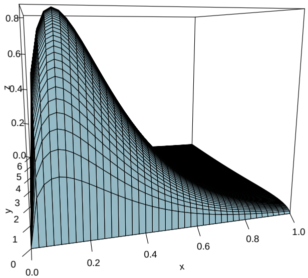


FIGURE: Density plot of the H-B distribution with $a = 1.5$, $b = 5$, $c = 0.5$ and $d = 1$.

Some properties of the H-B distribution

To make elicitation of the hyper-parameters, we need to investigate some properties of the H-B distribution. The second part of the bivariate, say Y can be marginalized to have Gamma distribution while the first part, say X, can not be marginalized to any closed form. However, the expectations and the variances of both components can be obtained.

Theorem

If $(X, Y) \sim H\text{-}B(a, b, c, d)$ *then*

(i) X *given* $Y = y$ *has a gamma distribution with parameters* (a, bd^y),

(ii) Y *has a gamma distribution with parameters* $\Big(a, \log(d^a/c)\Big)$.

Expectations and Variance of X and Y

The previous theorem allows us to compute the expectation and the variance of X and Y.

Theorem

Let $(X, Y) \sim H\text{-}B(a, b, c, d)$ *and denote* $k = \log(d^a/c)$*.*

$$E(Y) = a/k \qquad E(X) = \frac{a}{b}\left[\frac{k}{k + \log d}\right]^a$$

$$CV(Y) = a^{-1/2} \qquad CV(X) = a^{-1/2}\left[\frac{(k + \log d)}{\sqrt{k(k + 2\log d)}}\right]^a \tag{27}$$

Special case when $d = 1$

It is interesting to remark that in the case of $d = 1$, the two components of the bivariate random variable (X, Y) become independent and both of them have Gamma distribution with parameters (a, b) and $(a, \log(1/c))$ respectively. Note that in this case, the simple forms of the expectations and the variances can make the elicitation of the hyper-parameters become easy.

$$E(Y) = a/\log(1/c) \qquad E(X) = \frac{a}{b}$$
$$CV(Y) = a^{-1/2} \qquad CV(X) = a^{-1/2} \qquad (28)$$

This remark can be useful when we use the transformation data $(t_1, t_2, \ldots, t_n) \rightarrow (t_1/t_n, t_2/t_n, \ldots, 1)$

HB-distribution as natural conjugate prior for Bayesian analysis of the PLP

The following theorem show that the HB-distribution is a natural conjugate prior of the PLP.

Theorem

Let $\underline{t} = (t_1, \ldots, t_n)$ be the jump dates of the PLP with intensity $m(t) = \lambda\beta t^{\beta-1}$. The natural conjugate prior for Bayesian analysis of the PLP is H-B distribution with parameters (a, b, c, t_n) and the posterior is a H-B distribution with parameters $(a+n, b+1, cu, t_n)$.

HB-distribution as natural conjugate prior for Bayesian analysis of the PLP

The likelihood

$$L(\lambda, \beta) \propto \lambda^n \beta^n u^\beta \exp\left\{-\lambda t_n^\beta\right\}. \tag{29}$$

If the bivariate random variable $(\lambda, \beta) \sim \text{H-B}(a, b, c, t_n)$ then it has the joint density by the form

$$\pi(\lambda, \beta) \propto \lambda^{a-1} \beta^{a-1} c^\beta \exp\left\{-b\lambda t_n^\beta\right\}. \tag{30}$$

The posterior is the product of the likelihood and the prior

$$\begin{aligned}\pi(\lambda, \beta \mid \underline{t}) &\propto \pi(\lambda, \beta) \times L(\lambda, \beta) \\ &\propto \lambda^{a+n-1} \beta^{a+n-1} (cu)^\beta \exp\left\{-(b+1)\lambda t_n^\beta\right\}. \end{aligned} \tag{31}$$

It indicates that $(\lambda, \beta) \mid \underline{t} \sim \text{H-B}(a+n, b+1, cu, t_n)$.

Bayesian Estimators of α and β

Assuming a quadratic loss, the Bayesian estimators are the expectation of the posterior distributions. The Bayesian Estimator of parameter α and β are

$$\hat{\beta}_{Bayes3} = \frac{a+n}{k'+v}, \tag{32}$$

$$\hat{\lambda}_{Bayes3} = \left(\frac{a+n}{b+1}\right) \cdot \left(\frac{k'+v}{k'+v+\log(t_n)}\right)^{a+n} \tag{33}$$

where $k' = \log(t_n^a/c)$.

Bayesian Estimators as a convex combination of the MLEs and the expectation of the prior distribution

Recall that

$$E(\beta) = a/k', \qquad E(\lambda \mid \beta) = \frac{a}{bt_n^{\beta}}$$
$$\hat{\beta}_{MLE} = n/v, \qquad \hat{\lambda}_{MLE} = n/t_n^{\hat{\beta}_{MLE}} \tag{34}$$

We find that $\hat{\beta}_{Bayes3}$ can be expressed as a convex combination of the MLE and the expectation of the prior distribution

$$\hat{\beta}_{Bayes3} = p\hat{\beta}_{MLE} + (1-p)E(\beta) \tag{35}$$

where $p = \frac{v}{k'+v}$.

Bayesian Estimators as a convex combination of the MLEs and the expectation of the prior distribution

The relationship between $\hat{\lambda}_{Bayes3}$ and $\hat{\lambda}_{MLE}$ can be obtained approximately as following

$$\hat{\lambda}_{Bayes3} = \left(\frac{a+n}{b+1}\right) \cdot \left(\frac{1}{1+\log(t_n^{\hat{\beta}_{Bayes3}})/(a+n)}\right)^{a+n}$$
$$\approx \left(\frac{a+n}{b+1}\right) . \exp\left\{-\log(t_n^{\hat{\beta}_{Bayes3}})\right\} = \left(\frac{a+n}{b+1}\right) . \frac{n}{t_n^{\hat{\beta}_{Bayes3}}} \qquad (36)$$

Bayesian Estimators as a convex combination of the MLEs and the expectation of the prior distribution

Therefore, $\hat{\lambda}_{Bayes3}$ approximates a convex combination of the MLE and the prior expectation of λ given β

$$\hat{\lambda}_{Bayes3} = q\hat{\lambda}_{MLE} + (1-q)E(\lambda \mid \beta) \tag{37}$$

where $q = \frac{1}{b+1}$. These relationship will be very important to make elicitation on hyper-parameters.

Elicitation on hyper-parameters

Based on Bayesian Estimators of parameters α and β, we propose a strategy to integrate prior information on picking up values for hyper-parameters a, b, c. Suppose that we have prior information of the expectation and the coefficient of variation of β, say $E(\beta) = m$ and $CV(\beta) = r$, then the value for a and c can be easily obtained

$$a = \frac{1}{r^2}, c = \frac{t_n^a}{e^{a/m}} \tag{38}$$

Now we know the value of $p = \frac{v}{k'+v}$. We can find the value of b by setting $p = q$ or $p = 1 - q$ to obtain

$$b = \frac{1}{p} - 1. \tag{39}$$

Special case when $d = 1$

Recall that, if $(X, Y) \sim$ H-B$(a, b, c, 1)$ then X and Y are independent, X has Gamma distribution with parameters (a, b) and X has Gamma distribution with parameters $(a, \log(1/c))$. In last section we see that $(\lambda, \beta) \sim$ H-B(a, b, c, t_n) is the natural conjugate prior for PLP. We can transform the observation $(t_1, t_2, \ldots, t_{n-1}, t_n)$ to $(s_1, s_2, \ldots, s_{n-1}, s_n)$ where $s_i = t_i/t_n$.

Transformed data

The following theorem show how the PLP change with the transformed data.

Theorem

Let $(t_1, t_2, \ldots, t_{n-1}, t_n)$ be date of events of the Non-homogeneous Poisson Process $\{N(t), t > 0\}$ with intensity $m(t)$. For the positive constant $\gamma > 0$ and the transformation $s_i = t_i/\gamma; i = 1, \ldots, n$, the collection $(s_1, s_2, \ldots, s_{n-1}, s_n)$ is then date of events of the other Non-homogeneous Poisson Process $\{N^*(s), s > 0\}$ with intensity $m^*(s) = \gamma.m(\gamma s)$.

Transformed data

Assume that we have the observation $(t_1, t_2, \ldots, t_{n-1}, t_n)$ from the Power Law Process $\{N(t), t > 0\}$ with intensity $m(t) = \beta/\alpha(t/\alpha)^{\beta-1}$. Applying the transformation $s = t/t_n$, we have the observation $(s_1, s_2, \ldots, s_{n-1}, 1)$ from the Power Law Process $\{N^*(s), s > 0\}$ having intensity $m^*(s) = (t_n/\alpha)^\beta \beta s^{\beta-1}$. From now on, we can conduct Bayesian analysis for the Power Law Process $\{N^*(s), s > 0\}$.

Re-parametrization

Because the intensity has the form $m^*(s) = (t_n/\alpha)^\beta \beta s^{\beta-1}$, it is reasonable to re-parametrize $\eta = (t_n/\alpha)^\beta$. The intensity and the mean function become $m^*(s) = \eta\beta s^{\beta-1}$, $M^*(s) = \eta s^\beta$ respectively. The likelihood is

$$\begin{aligned} L(\eta,\beta) &= \left[\prod_{i=1}^{n} m^*(s_i)\right].\exp\left\{-M^*(s_n)\right\} \\ &= \eta^n\beta^n\left[\prod_{i=1}^{n} s_i\right]^{\beta-1} e^{-\eta} \\ &\propto \eta^n e^{-\eta}\times\beta^n e^{-\log(1/w)\beta} \end{aligned} \tag{40}$$

where $w = \prod_{i=1}^{n} s_i < 1$.

Re-parametrization

The natural conjugate prior joint distribution for (η, β) turns out to be

$$(\eta, \beta) \sim Gamma(a_\eta, b_\eta) \times Gamma(a_\beta, b_\beta). \tag{41}$$

The posterior is then

$$(\eta, \beta \mid \underline{s}) \sim Gamma(a_\eta + n, b_\eta + 1) \times Gamma(a_\beta + n, b_\beta + \log(1/w)), \tag{42}$$

where $\underline{s} = \{s_1, \ldots, s_{n-1}, 1\}$.

Re-parametrization

It means that η and β are independent and each parameter is distributed by Gamma law

$$\eta \mid \underline{s} \sim Gamma(a_\eta + n, b_\eta + 1) \tag{43}$$

$$\beta \mid \underline{s} \sim Gamma(a_\beta + n, b_\beta + \log(1/w)). \tag{44}$$

Note that

$$1/w = 1/\prod_{i=1}^{n} s_i = \prod_{i=1}^{n} (t_n/t_i) \tag{45}$$

therefore

$$\log(1/w) = v = n/\hat{\beta}_{MLE}. \tag{46}$$

Re-parametrization

Here we can see again exact the same result of Olivera et al. (2012). Therefore, applying H-B distribution as the natural conjugate prior is more general than independent conjugate prior by Gamma distribution. Using the method of re-parametrization is quit useful for Bayesian analysis. Note that Maximum Likelihood Estimator is invariant under re-parametrization but Bayes estimator is not. Let's look at the summary table before we go to the application section.

Summary table

Table: Table of parametric estimation for PLP

Maximum Likelihood Estimation	
MLE for β	MLE for α, θ, λ
$\hat{\beta}_{MLE} = n/v$	$\hat{\alpha}_{MLE} = t_n/n^{1/\hat{\beta}_{MLE}}$
	re-parametrize $\theta = (t_n/\alpha)^{\beta}$
$u = \prod_{i=1}^{n} t_i$	$\hat{\theta}_{MLE} = n$
	re-parametrize $\lambda = 1/\alpha^{\beta}$
$v = \sum_{i=1}^{n} \log(t_n/t_i)$	$\hat{\lambda}_{MLE} = n/t_n^{\hat{\beta}_{MLE}}$

Summary table

TABLE: Table of parametric estimation for PLP

Non-informative Prior	
Prior distribution	
$\pi(\alpha,\beta) \propto (\alpha,\beta)^{-1}$	
Posterior distribution	
$\pi(\alpha,\beta \mid \underline{t}) \propto K\alpha^{-(n\beta+1)}\beta^{n-1}u^{\beta}e^{-(t_n/\alpha)^{\beta}}$	$K = v^{n-1}/(\Gamma(n)\Gamma(n-1))$
Bayes Estimate for β	Bayes Estimate for α
$\hat{\beta}_{Bayes1} = (n-1)/v$	$\hat{\alpha}_{Bayes1}$ has no close form

Summary table

TABLE: Table of parametric estimation for PLP

Independent conjugate prior with two Gamma distributions	
Prior distribution	
$\pi(\theta,\beta) \sim Gamma(a_\theta, b_\theta) \times Gamma(a_\beta, b_\beta)$	
Posterior distribution	
$\pi(\theta,\beta \mid \underline{t}) \sim Gamma(a_\theta + n, b_\theta + 1)$ $\times Gamma(a_\beta + n, b_\beta + v)$	
Bayes Estimate for β	Bayes Estimate for θ
$\hat{\beta}_{Bayes2} = (a_\beta + n)/(b_\beta + v)$	$\hat{\theta}_{Bayes2} = (a_\theta + n)/(b_\theta + 1)$
Convex combination	Convex combination
$\hat{\beta}_{Bayes2} = p\hat{\beta}_{MLE} + (1-p)E(\beta)$	$\hat{\theta}_{Bayes2} = q\hat{\theta}_{MLE} + (1-q)E(\theta)$
$p = v/(b_\beta + v)$	$q = 1/(b_theta + 1)$

Summary table

TABLE: Table of parametric estimation for PLP

Natural conjugate prior with H-B distribution Prior distribution $\pi(\lambda, \beta) \sim H - B(a, b, c, t_n)$ Posterior distribution $\pi(\lambda, \beta \mid \underline{t}) \sim H - B(a+n, b+1, cu, t_n)$	
Bayes Estimate for β $\hat{\beta}_{Bayes3} = (a+n)/(k+v)$ $k = \log(t_n^a/c)$ Convex combination $\hat{\beta}_{Bayes3} = p\hat{\beta}_{MLE} + (1-p)E(\beta)$ $p = v/(k+v)$	Bayes Estimate for λ $\hat{\lambda}_{Bayes3} = (a+n)/(b+1).$ $[(k+v)/(k+v+\log(t_n))]^{a+n}$ Convex combination $\hat{\lambda}_{Bayes3} = q\hat{\lambda}_{MLE} + (1-q)E(\lambda \mid \beta)$ $q = 1/(b+1)$

Simulation data

In order to investigate the behaviour of the H-B distribution as natural conjugate prior, we make a comparison between Bayesian estimation and MLE relying on simulated data from PLP. We simulating three sets of data varying from small size $n = 20$ to medium size $n = 150$ and then to very large size $n = 2000$. The three data sets generated by a PLP with true parameters $\beta = 1.38, \lambda = 0.008$. The case of small size is in favour to show the advantage of Bayesian approach.

Prior mean and variability

For the elicitation strategy, three different values of prior means for β are investigated

(1) Prior mean underestimates the true value.

(2) Prior mean is relatively close to the true value.

(3) Prior mean overestimates the true value.

In each scenario of prior mean guess m, we use three different degree of variability $r = 0.3$ (corresponding to relatively small variability), $r = 0.6$ (corresponding to relatively moderate variability), and $r = 0.9$ (corresponding to relatively large variability).

Bayesian estimation for simulated data from a PLP with input parameter values $\beta = 1.38$ and $\lambda = 0.0008$

Sample-size	Prior guess		Bayes estimates	
n	m	r	$\hat{\beta}_{Bayes3}$	$\hat{\lambda}_{Bayes3}$
10	0,90	0,27	1,1338	0,006949
		0,54	1,4009	0,012574
		0,81	1,5412	0,014838
	1,40	0,42	1,4973	0,002126
		0,84	1,6157	0,007978
		1,26	1,6750	0,011449
	2,10	0,63	1,8461	0,000892
		1,26	1,7543	0,006538
		1,89	1,7318	0,010739
	MLE		1.4343	0.001604

Bayesian estimation for simulated data from a PLP with input parameter values $\beta = 1.38$ and $\lambda = 0.0008$

Sample-size	Prior guess		Bayes estimates	
n	m	r	$\hat{\beta}_{Bayes3}$	$\hat{\lambda}_{Bayes3}$
150	0,90	0,27	1,3545	0,001994
		0,54	1,3936	0,001714
		0,81	1,4016	0,001662
	1,40	0,42	1,4114	0,001363
		0,84	1,4124	0,001506
		1,26	1,4127	0,001535
	2,10	0,63	1,4464	0,001085
		1,26	1,4226	0,001411
		1,89	1,4180	0,001484
	MLE		1.3995	0.001082

Bayesian estimation for simulated data from a PLP with input parameter values $\beta = 1.38$ and $\lambda = 0.0008$

Sample-size	Prior guess		Bayes estimates	
n	m	r	$\hat{\beta}_{Bayes3}$	$\hat{\lambda}_{Bayes3}$
2000	0,90	0,27	1,3848	0,000904
		0,54	1,3879	0,000881
		0,81	1,3885	0,000877
	1,40	0,42	1,3855	0,000899
		0,84	1,3855	0,000904
		1,26	1,3854	0,000905
	2,10	0,63	1,3906	0,000854
		1,26	1,3887	0,000875
		1,89	1,3883	0,000879
	MLE		1.3803	0.000834

Commentary

Table 1 describes the result of Bayesian Estimation based on the simulated data from a PLP with input parameter values $\beta = 1.38$ and $\lambda = 0.0008$. We take $m = 0.9$ for underestimated guess of prior mean, $m = 1.4$ for relative accurate guess of prior mean and $m = 2.1$ for overestimated guess of prior mean.

In case of large sample size, it is not surprising that Bayesian estimates are relatively close to MLEs and both are close to the true values of λ and β. That fact holds no matter how the prior mean underestimates or overestimates the true value of the parameters. For small and medium size, one can see that the underestimating scenario is more accurate than the two other scenarios. More detail, the Bayesian estimates seem to increase with respect to the variability. In case of medium sample size, the underestimated prior guess with moderate variability $r = 0.6$ Bayesian estimators produce more accurate result than MLEs but in small sample size case, the MLEs tend to perform better.

Concluding remarks

In this work we introduce a new distribution : the H-B distribution. This distribution is a natural conjugate prior to make Bayesian inference on the PLP. More investigations concerning the properties of this distribution need to be carried out. In particular a better understanding of the properties will be helpful to elicit prior hyper-parameters. Our strategy is quit easy to implement, relying on expert guessing. The simulation result shows that the choice of the elicitation strategy is very sensitive. More need to be done in order to improve the accuracy of the estimates. Other strategies should be investigated. We are working in this direction in the present time.

Thank you for your attention !

For references of Power Law Process, please see (**?**), (**?**).