Computational Finance Project

Sarthak Chowdhuri (MB1623)
Arju Jana (MB1604)

ISI Kolkata
MStat 2nd Yr.

April 23,2018

Problem Statement

Download historical daily closing prices available in [a] for the past 10 years of the constituents of NSE Nifty 50 [b]. Estimate the indices and proportion of a mixture model for the tail index. For a single population, the estimation can be done using [c].

- Historical data [a]
- Nifty stock list [b]
- Extreme Fit in R [c]

	Company.Name <fctr>	Industry <fctr>	Symbol <fctr>	Series <fctr>	ISIN.Code <fctr>
1	Adani Ports and Special Economic Zone Ltd.	SERVICES	ADANIPORTS	EQ	INE742F01042
2	Asian Paints Ltd.	CONSUMER GOODS	ASIANPAINT	EQ	INE021A01026
3	Axis Bank Ltd.	FINANCIAL SERVICES	AXISBANK	EQ	INE238A01034
4	Bajaj Auto Ltd.	AUTOMOBILE	BAJAJ-AUTO	EQ	INE917I01010
5	Bajaj Finance Ltd.	FINANCIAL SERVICES	BAJFINANCE	EQ	INE296A01024
6	Bajaj Finserv Ltd.	FINANCIAL SERVICES	BAJAJFINSV	EQ	INE918I01018

Figure 1: Nifty50 List

Nifty50

AUTOMOBILE	CEMENT & CEMENT PRODUCTS	CONSTRUCTION
6	2	1
CONSUMER GOODS	ENERGY	FERTILISERS & PESTICIDES
4	8	1
FINANCIAL SERVICES	IT	MEDIA & ENTERTAINMENT
11	5	1
METALS	PHARMA	SERVICES
4	4	1
TELECOM		
2		

Figure 2: Industry wise classification

We look at their daily closing stock prices for the last 10 years. For each constituent, we are to estimate the indices and proportion of a mixture model for the tail index. Let us work with one of them, namely SBI.

SBI Historical Data

We have downlaoded the data from Investing.com source which was given to us.

	ï..Date	Price	Open	High	Low	Vol.	Change..
1	Jan 01, 2018	307.10	310.6	312.75	306.30	12.18M	-0.90%
2	Dec 29, 2017	309.90	310.0	312.00	309.05	11.94M	0.49%
3	Dec 28, 2017	308.40	315.3	316.50	307.65	20.35M	-2.05%
4	Dec 27, 2017	314.85	316.5	320.30	313.05	14.28M	-0.73%
5	Dec 26, 2017	317.15	318.6	319.95	316.30	9.33M	-0.84%
6	Dec 22, 2017	319.85	317.1	323.85	316.50	13.82M	0.98%

Figure 3: Stock Price (SBI)

We will look at the daily closing prices, and then consider the log returns. Let us look at a few empirical things.

Parameter	Estimate
Mean	0.0001258867
Std Dev.	0.02378821
Skewness	0.6674823
Kurtosis	10.70131

Log Return Density

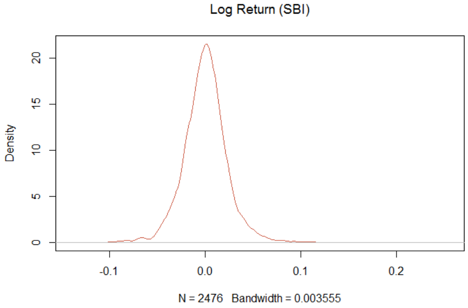


Figure 4: Log Return (SBI)

Using Single General Pareto Distributions
The pdf is given by

$$f(x) = \frac{1}{\lambda}[1 + \frac{1}{\alpha}(\frac{x-u}{\lambda})]^{-(\alpha+1)} \tag{1}$$

with $x \geqq u, \alpha > 0, \lambda > 0$.

Parameter Estimation

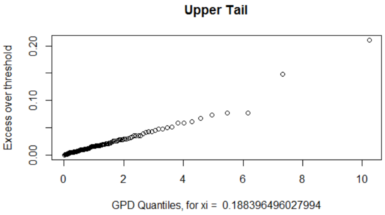
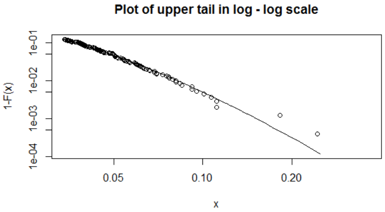


Figure 5: Mean Excess Plot (SBI)



Parameter Estimation

We can further be more careful while choosing the estimated parameters, by looking for stability from the below plot.

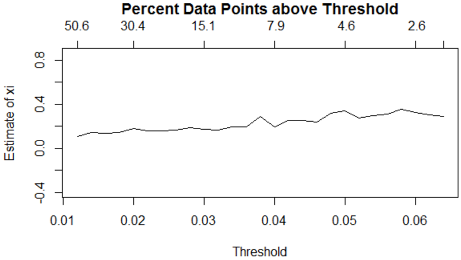


Figure 7: Shape Parameter ξ with Threshold (SBI)

Pareto Mixture

Using Mixture of General Pareto Distributions
We are to implement an EM Algorithm to model the tail as a mixture of 2 Pareto distributions estimating the shapes, scales and proportions.
Here is the theory for the same.
$f_1(x) = \frac{1}{\lambda_1}[1 + \frac{1}{\alpha_1}(\frac{x-x_1}{\lambda_1})]^{-(\alpha_1+1)} I(x > x_1)$
$f_2(x) = \frac{1}{\lambda_2}[1 + \frac{1}{\alpha_2}(\frac{x-x_2}{\lambda_2})]^{-(\alpha_2+1)} I(x > x_2)$
In specific we consider $x_1 = 0$ and $x_2 = x_i$.

We define,

$$Z_i = \begin{cases} 1 & X_i \sim f \\ 0 & X_i \sim g \end{cases}$$

$P(Z_i = 1) = \pi_1$,
So,

$$\begin{aligned} E[Z_i|X = x] &= P(Z_i = 1|X = x) \\ &= \frac{\pi_1 f_1(x|\alpha_1, \lambda_1)}{\pi_1 f_1(x|\alpha_1, \lambda_1) + (1 - \pi_1) f_2(x|\alpha_2, \lambda_2)} \end{aligned} \tag{2}$$

Expectation-Maximization

Consider the likelihood for $\forall x > 0$,

$$\textit{Likelihood } L = \prod_{1}^{n} (\pi_1 f_1(X_i))^{Z_i} (\pi_2 f_2(X_i))^{1-Z_i} \tag{3}$$

$$\Rightarrow \log L = \sum_{i=1}^{n} [Z_i(\log \pi_1 - \log \lambda_1 - (\alpha_1 + 1)\log(\frac{x - x_1}{\lambda_1}) + I(x > x_1))$$
$$+(1 - Z_i)(\log \pi_2 - \log \lambda_2 - (\alpha_2 + 1)\log(\tfrac{x-x_2}{\lambda_2}) + I(x > x_2))]$$

E-Step

$$E(\log\ L) = \sum_{i=1}^{n} [\hat{Z}_i(\log \pi_1 - \log \lambda_1 - (\alpha_1 + 1)\log(\frac{x - x_1}{\lambda_1}) + I(x > x_1))$$
$$+(1 - \hat{Z}_i)(\log \pi_2 - \log \lambda_2 - (\alpha_2 + 1)\log(\tfrac{x-x_2}{\lambda_2}) + I(x > x_2))]$$

Expectation-Maximization

M-Step

$$\frac{\partial \log E(\log L|x_1,x_2,..x_n)}{\partial \pi_1} = 0$$

$$\Rightarrow \frac{1}{\pi_1}\sum_{i=1}^{n}\hat{Z}_i - \frac{1}{1-\pi_1}\sum_{i=1}^{n}(1-\hat{Z}_i) = 0$$

$$\Longrightarrow \hat{\pi_1}^{(1)} = \frac{1}{n}\sum_{i=1}^{n}\hat{Z}_i^{(0)}$$

$$\frac{\partial \log E(\log L|x_1,x_2,..x_n)}{\partial \alpha_1} = 0$$

$$\Rightarrow (\alpha_1+1)\sum_{i=1}^{n}\frac{\hat{Z}_i}{1+\frac{x_i-x_1}{\alpha_1\lambda_1}}\frac{x_i-x_1}{\alpha_1^2\lambda_1} = \hat{Z}_i\log(1+\frac{x-x_1}{\alpha_1\lambda_1})$$

$$\Longrightarrow \hat{\alpha_1}^{(1)} = \left[\frac{\hat{Z}_i^{(0)}\log(1+\frac{x_i-x_1}{\alpha_1^{(0)}\lambda_1^{(0)}})}{\sum_{i=1}^{n}\frac{\hat{Z}_i^{(0)}}{1+\frac{x_i-x_1}{\alpha_1^{(0)}\lambda_1^{(0)}}}(\frac{x_i-x_1}{\alpha_1^{(0)}\lambda_1^{(0)}})} - 1\right]^{-1}$$

Expectation-Maximization

$$\frac{\partial \log E(\log L|x_1,x_2,..x_n)}{\partial \alpha_2} = 0$$

$$\Rightarrow (\alpha_2+1)\sum_{i=1}^{n}\frac{1-\hat{Z}_i}{1+\frac{x_i-x_1}{\alpha_2\lambda_2}}\frac{x_i-x_1}{\alpha_2^2\lambda_2} = (1-\hat{Z}_i)\log(1+\frac{x_i-x_1}{\alpha_2\lambda_2})$$

$$\Longrightarrow \hat{\alpha_2}^{(1)} = \left[\frac{(1-\hat{Z_i}^{(0)})\log(1+\frac{x_i-x_1}{\alpha_2^{(0)}\lambda_2^{(0)}})}{\sum_{i=1}^{n}\frac{1-\hat{Z_i}^{(0)}}{1+\frac{x_i-x_1}{\alpha_2^{(0)}\lambda_2^{(0)}}}(\frac{x_i-x_1}{\alpha_2^{(0)}\lambda_2^{(0)}})} - 1\right]^{-1}$$

$$\frac{\partial \log E(\log L|x_1,x_2,..x_n)}{\partial \lambda_1} = 0$$

$$\Rightarrow (\alpha_1+1)\sum_{i=1}^{n}\frac{\hat{Z}_i}{1+\frac{x_i-x_1}{\alpha_1\lambda_1}}\frac{x_i-x_1}{\alpha_1\lambda_1^2} = \frac{1}{\lambda_1}\sum_{i=1}^{n}\hat{Z}_i$$

$$\Longrightarrow \hat{\lambda_1}^{(1)} = \frac{(\alpha_1^{(0)}+1)\sum_{i=1}^{n}\frac{\hat{Z_i}^{(0)}}{1+\frac{x_i-x_1}{\alpha_1^{(0)}\lambda_1^{(0)}}}\frac{x_i-x_1}{\alpha_1^{(0)}\lambda_1^{(0)}}}{\sum_{i=1}^{n}\hat{Z_i}^{(0)}}$$

Expectation-Maximization

Similarly we have, $\Longrightarrow \hat{\lambda_2}^{(1)} = \dfrac{(\alpha_2^{(0)}+1) \displaystyle\sum_{i=1}^{n} \frac{1-\hat{Z_i}^{(0)}}{1+\frac{x_i - x_1}{\alpha_2^{(0)}\lambda_2^{(0)}}} \frac{x_i - x_1}{\alpha_2^{(0)}\lambda_2^{(0)}}}{\displaystyle\sum_{i=1}^{n}(1-\hat{Z_i}^{(0)})}$

few points!

- However using unknown locations, the EM algorithm will not work as they will never be updated. Here is the reference paper for the same in ResearchGate website. "On Maximum Likelihood Estimation of a Pareto Mixture, January 2103, by Marco Bee, Giuseppe Espa, Roberto Benedetti".
  So we take them as known values from the single distribution fit. Then we carry on applying the EM Algorithm.
- It is interesting to see that one of the Pareto Distribution of the mixture is highly dominant and its parameters are very close to what we get when we used a single GPD to estimate.Similarly the lower tail and also for the other stocks.

Results

In the same way we analyze for the other stocks. We summarize the results for a few below:

Stock	Single GPD shape	Single GPD scale	Mixture GPD dominant shape	Mixture GPD dominant
SBI	5.610924	0.01452079	5.610822	0.01452078
COAL	8.578711	0.01076159	8.578711	0.01076242
CNTY	7.686716	0.01745051	7.685447	0.01744987
DSTV	4.615578	0.01878393	4.615227	0.01878339
EICH	3.364470	0.01531268	3.364226	0.01531208
GLEN	4.584767	0.01627197	4.584292	0.01627135
HDBK	5.892666	0.01194083	5.89262	0.01194082
HDFC	4.519679	0.01381828	4.519149	0.01381763
HLL	5.472644	0.01028556	5.571171	0.01031392
HPCL	7.886756	0.01553146	7.885558	0.01553085
ICBK	24.27835	0.02093861	24.27837	0.02093860
INFY	4.593573	0.01158688	4.593227	0.01158677
ITC	8.905622	0.01090664	8.904653	0.01090651
LICH	6.932689	0.01904173	6.931799	0.01904114
LUPN	25.29333	0.01179709	25.29339	0.01179709
MAHM	2.966781	0.01128694	2.972521	0.01129308
ONGC	5.952683	0.01258339	5.952454	0.01258333

Figure 8: Results for other Stocks

Results

Stock	Single GPD shape	Single GPD scale	Mixture GPD dominant shape	Mixture GPD dominant
RANB	4.165468	0.01594564	4.161342	0.01594299
REDY	24.15001	0.01187990	24.13921	0.01187973
RELI	4.157552	0.01030483	4.255115	0.01034989
TAMO	8.378704	0.01706061	8.377153	0.01705996
TCS	7.51643	0.01460201	7.515213	0.01460135
TEML	3.565791	0.01597966	3.566564	0.01598011
TTEX	14.38224	0.02580105	14.38006	0.02580058
UPLL	4.380475	0.01540045	4.383597	0.01540248
WCKH	3.993936	0.02069185	3.993621	0.02069126
YESB	3.501047	0.01621788	3.512563	0.01623037
BAJA	5.194496	0.01067211	5.193866	0.01067194
BFRG	5.87405	0.01652911	5.874944	0.01652997
BOB	2.898630	0.01336198	2.898461	0.01336140
BPCL	10.75635	0.01472858	10.75326	0.01472785
APSE	7.899816	0.01836482	7.90111	0.01836565
ARBN	13.67091	0.01807263	13.6689	0.01807249
ASOK	5.610924	0.01452079	5.610822	0.01452078
AXBK	15.44735	0.01783171	15.45257	0.01783257

Figure 9: Results for other Stocks

Bayesian Framework[1]
Pareto and it's Estimates

- Let Y is distributed according to a Pareto law.
  It's density $f(y|\alpha, y_m) = \alpha(y_m)^{\alpha} y^{-(\alpha+1)} I(y > y_m), \qquad y_m > 0, \alpha > 0$

  The cumulative distribution is $F(y) = (1 - y_m^{\alpha} y^{-\alpha}) I(y > y_m)$
- From the likelihood equation we get the two sufficient statistics as $Min(y)$ and $\sum_{i=1}^{n} \log(y_i/y_m)$. Whose classical estimates are obtained by taking $\hat{y_m} = Min(y)$ and $\hat{\alpha} = n/\sum_{i=1}^{n} \log(y_i/y_m)$.

$$L(y|\alpha, y_m) = \alpha^n exp[-(\alpha+1)\sum_{i=1}^{n} \log(y_i) + \alpha n \log(y_m)] I(y_{(1)} > y_m) \tag{4}$$

[1] Ndoye, A.A. and Lubrano, M., 2014, September. Tournaments and Superstar models: A Mixture of two Pareto distributions. In Economic Well-Being and Inequality: Papers from the Fifth ECINEQ Meeting (pp. 449-479). Emerald Group Publishing Limited.

Power Distribution

- X is said to have a power function distribution if its probability density function is defined as,

$$p(x) = \alpha x_m^{-\alpha} x^{\alpha-1} I(x < x_m) \tag{5}$$

where $\alpha > 0 and x_m > 0$.

- Cumulative distribution function, $F(x) = x_m^{-\alpha} x^{\alpha} I(x < x_m)$
- Two sufficient statistics are provided by $Max(y) and \sum_{i=1}^{n} \log(y_i/y_m)$

note: If x has a power function distribution in (α, x_m), then $y = 1/x$ is distributed according to a $Pareto(\alpha, y_m)$ where $y_m = 1/x_m$.

Pareto Mixture

Mixture of two pareto with different shape and scale parameters,
$f(y|\alpha_1, \alpha_2, y_{m1}, y_{m2}, p) =$

$$p\alpha_1 y_{m1}^{\alpha_1} y^{-(\alpha_1+1)} I(y > y_{m1}) + (1-p)\alpha_2 y_{m2}^{\alpha_2} y^{-(\alpha_2+1)} I(y > y_{m2}) \qquad (6)$$

- The two components have a different support, so it is natural to assume for instance that $y_{m2} > y_{m1}$. In this framework, the first member is concerned with observations greater than y_{m1} while the second component corresponds to observations greater than y_{m2}. So, any observation y_i such that $y_{m1} < y_i < y_{m2}$ belongs to the first regime with probability 1 and not with probability p.
- Since the usual EM algorithm does not work for estimating the five parameters, we present Gibbs Sampling technique to estimate the parameters.

Conjugate Priors

The priors we have used for this Bayesian framework,

- Prior on α : $Gamma(\alpha_0, \nu_0)$

  Posterior: $Gamma(\alpha_0 + \sum_{i=1}^{n} \log(y_i/y_m), \nu_0 + n)$
- Prior on y_m $Power(\gamma_0, y_{m0})$
  Posterior : $Power(\gamma_0 + n\alpha, Max(Min(y_i), y_{m0}))$
- Prior for p : $Beta(n_{01}, n_{02})$
  Posterior: $Beta(n_{01} + n_1, n_{02} + n_2)$, where n_1 and n_2 are the number of observations that conditionally on z fall into each regime.

Gibbs Sampler

- Fix a value for the total number of draws m, fix a value for y_{m2}, select a starting value for p, and compute the following starting values $y_{m1} = y_{(1)}, \alpha_1 = \hat{\alpha}(y_{m1}), \alpha_2 = \hat{\alpha}(y_{m2})$.
- Start the loop on j, the Gibbs iterations.
- Determine the observations $y_{1s}|y < y_{m2}$ that belong for sure to the first regime for a given draw of y_{m2}.Determine the remaining observations $y_{12}|y > y_{m2}$.
- For the remaining observations y_{12}, simulate the sample allocation $z^{(j)}$ where each element is drawn according to a $Binomial(z^{(i)}|p_i)$, with base probability $p_i = \frac{pf_1}{pf_1+(1-p)f_2}$.

Gibbs Sampler (Further Steps)

- Select the sub-sample separation y_{1r} and y_{2r} among the y_{12}.
- Form the first regime allocation $y_1^{(j)} = (y_{1s}, y_{1r})$, and the second regime allocation $y_2^{(j)} = y_{2r}^{(j)}$.
- Compute $n_1^{(j)} = n_{1s}^{(j)} + n_{1r}^{(j)}$ and $n_2^{(j)}$.
- Draw $p^{(j)} \sim Beta(n_1^{(j)} + n_{01}, n_2^{(j)} + n_{02})$.
- Draw $y_{m1}^{(j)} \sim Power(\gamma_1^{(0)} + n_1^{(j)}\alpha_1^{(j-1)}, Max(Min(y_1^{(j)}), y_{m01}))$
- Draw $y_{m2}^{(j)} \sim Power(\gamma_2^{(0)} + n_2^{(j)}\alpha_2^{(j-1)}, Max(Min(y_2^{(j)}), y_{m02}))$
- Draw $\alpha_1^{(j)} \sim Gamma(\alpha_1^{(0)} + \sum_{i=1}^{n_1^{(j)}} \log(y_1^{(j)}/y_{m1}^{(j)}), \nu_{01} + n_1)$
- Draw $\alpha_2^{(j)} \sim Gamma(\alpha_2^{(0)} + \sum_{i=1}^{n_2^{(j)}} \log(y_2^{(j)}/y_{m2}^{(j)}), \nu_{02} + n_2)$
- $j = j + 1$

Prior Value for y_{m2} (SBI)

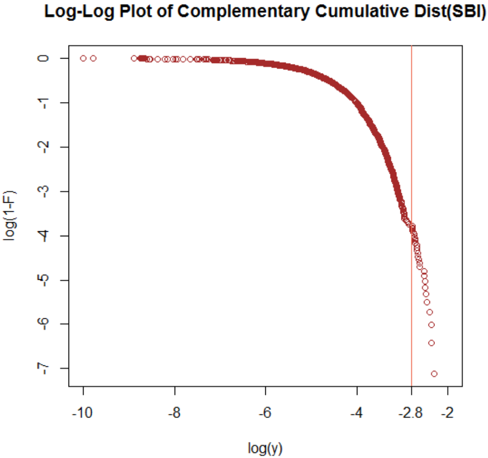


Figure 10: 1-Fn(x) vs log(y)(SBI)

Prior Value for y_{m2} (COAL)

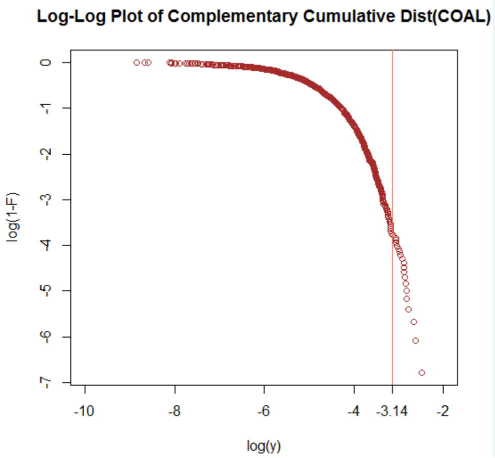


Figure 11: 1-Fn(x) vs log(y)(COAL)

Posterior Inference for Mixture Distribution

SBI

Parameter	Estimate
α_1	0.23148
α_2	0.9373483
y_{m1}	4.500561e-05
y_{m2}	0.006122063
p	0.3068412

COAL

Parameter	Estimate
α_1	0.2898847
α_2	1.45775
y_{m1}	0.0001407731
y_{m2}	0.009020065
p	0.5519838

Comparison: Cumulative Pareto Mixture Distribution with Empirical Distribution (SBI)

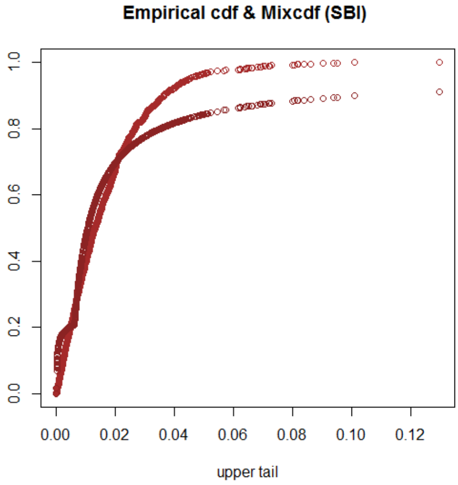


Figure 12: $F_n(x)$ and mixcdf (SBI)

Comparison: Cumulative Pareto Mixture Distribution with Empirical Distribution (COAL)

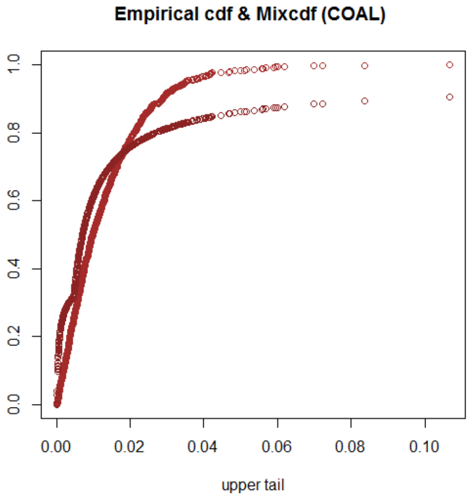


Figure 13: $F_n(x)$ and mixcdf (COAL)

few points!

- In Gibbs Sampling algorithm, the theory ensures that after a sufficiently large number of iterations, T, the set $(\alpha_1^{(j)}, \alpha_2^{(j)}, y_{m1}^{(j)}, y_{m2}^{(j)}, p^{(j)}) : j = T+1, \ldots, N$ can be seen as a random sample from the joint posterior distribution.
- This Bayesian inference for a mixture of Pareto is sensitive to the choice of prior information. We need to be careful on choosing the parameter's priors of the mixture.
- In Gibbs Sampler $\hat{\alpha}$ is computed using $\sum_{i=1}^{n} \log(y_i/y_m)$ where y_i is restricted to the sub-sample $y_i > y_m$. Same way we can also model for the log losses and for other stocks.

Thank You :-)