

Probabilistic Patterns Among Coronavirus Confirmed, Cured and Deaths in Thirty-two India's States/Territories

Ramalingam Shanmugam¹

¹ Honorary Professor of International Studies, School of Health Administration
Texas State University, San Marcos, TX 78666, USA
E-mail: rs25@txstate.edu

ABSTRACT

In this article, we propose, fit, and promote an incidence rate restricted Poisson distribution for the gathered number of coronavirus (1) confirmed, (2) cured, and (3) deaths across thirty two India's states/territories as of May 1, 2020. The restriction on the incidence rate in each of the three categories is quantitatively estimated and interpreted in terms of the attained administrative efficiency of preventive/treatment procedures by the government, using the public domain data in <http://mohfw.gov.in> webpage. The statistical conclusions based on the data analytic results demonstrate a valuable lesson to be procured to deal with any future pandemic episode.

Keywords: Incidence rate restricted Poisson distribution, index of dispersion, risk reduction, odds ratio, hazard rate, stability index

Mathematical Subject Classification: 60E05, 62-07, 62F03, 62P10

Computing Classification: C15, C46

1. INTRODUCTION

Whether the corona virus is a natural or human manufactured, it is a global pandemic, causing citizens in almost every nation on earth are locked inside their home with a social distancing to reduce, if not totally stop, the spread. The land, sea, and air transportations are shut down. The human mobility is astonishingly nonexistent. The schools, offices, and industries are closed, making the gross domestic productivity (GDP) and gross national productivity (GNP) to near zero level. The economic growth is so bleak for all nations with no exception. Not many nations have been successful to stop the spread of this ugly and fatal pandemic. India is one of the nations who have successfully slowed down, if not completely arresting the spread of the coronavirus by adapting a social distancing. Yet, India had to do a massive diagnostic checking of its citizens to check whether they have contracted the corona virus. A count of *coronavirus confirmed* cases was made available to the public's view as India enjoys a full freedom communication. The government concurrently transformed the confirmed individuals to designated hospitals and/or primary medical care centers. A count of how many of those became cured along with a number of deaths due to corona virus were made available to the public. These numbers were periodically compiled and displayed in a web page <http://mohfw.gov.in>. The numbers in the webpage are classified as: (1) Y_1 = confirmed coronavirus cases, (2) Y_2 = cured cases, and (3) Y_3 =

deaths due to coronavirus for thirty-two India's states/territories. See the data in Table 1 below as of May 1, 2020. The data seem to change on daily basis. The research questions that are postulated and then answered via probability model and statistical reasonings with the help of the collected data are how effective were the public testing for the prevalence of the corona virus, the treatment in the hospitals/clinics, and the pattern of deaths across the states/territories. Is the effectiveness similar or different in surrounding states/territories? If they are different among the neighboring states, what might be causing the differences? Could the governments and/or medical professionals learn from the data analytics-based knowledge, follow the lead in the results to refine their administrative or policy approaches to pandemics like the corona virus were to occur in future? These are the research aims of this article.

Table 1. Coronavirus confirmed, cured, and deaths in India.

States/territories	$Y_1 = \text{Confirmed}$	$Y_2 = \text{Cured}$	$Y_3 = \text{Deaths}$
Andaman Nicobar	33	16	0
Andhra Pradesh	1525	441	33
Arunachal Pradesh	1	1	0
Assam	43	32	1
Bihar	471	98	3
Chandigarh	88	17	0
Chhattisgarh	43	36	0
Delhi	3738	1167	61
Goa	7	7	0
Gujarat	4721	735	236
Haryana	360	227	4
Himachal Pradesh	40	30	1
Jammu Kashmir	639	247	8
Jharkhand	111	20	3
Karnataka	598	255	25
Kerala	498	392	4
Ladakh	22	17	0
Madhya Pradesh	2719	524	145
Maharashtra	11506	1879	485
Manipur	2	2	0
Meghalaya	12	0	1
Mizoram	1	0	0
Odisha	154	55	1
Puducherry	8	5	0
Punjab	772	112	20
Rajasthan	2666	1116	62
Tamil Nadu	2526	1312	28
Telangana	1057	441	26
Tripura	2	2	0
Uttar Pradesh	2455	656	43
Uttarakhand	58	37	0
West Bengal	795	139	33
Average, $\bar{y}_i$	1177	313	38

Variance, $s_{y_i}^2$	5074202.2	217381	9012.9
Incidence rate, $\hat{\lambda}_i =$	17.93	11.88	2.48
Restriction rate, $\hat{\beta}_i =$	18.20	12.34	2.66
$(1 - \frac{\hat{\lambda}_i}{\hat{\beta}_i})^2 =$	0.0002 (measure of effective social distancing index)	0.0014 (lack of imbalance curability index)	0.0042 (measure of unreliability to safety index)
Gap: $\delta_i = \beta_i - \lambda_i =$	0.27	0.46	0.17
p-value for $H_0 : \beta_i = \infty$	0.00001	0.00001	0.00001
power for $H_1 : \beta_{i,1} = 25$	0.37	0.12	0.99
OR_i in (1.5)	1	1	1
Hazard rate reduction in (1.6)	7.69	5.53	1.98
Vulnerability index in (1.7)	0.99	0.99	0.99
Stability index in (1.8)	0.00001	0.0001	0.0001
Tail value risk reduction in (1.9)	75602	1542	4.53

2. BACKGROUND INFORMATION BEFORE PROPOSING AN UNDERLYING MODEL

From the data in Table 1, an average of 1,177 per state has contracted coronavirus and it is much smaller than what it is Europe, Americas, or Southeast Asian countries. About 26% and 3% of them per state were cured in hospital or clinics and died. The data exhibit a high amount of variance compared to the average implying the prevalence of corona virus incidences is heterogeneous across all three categories. Such patterns are clues for an existence of incidence rate restricted Poissonness in the data. The data are counts and not normally distributed according to their probability-probability (P-P) plots. The data are likely to be a Poisson type. But the characteristic property of the Poisson distribution seems to be not validated by the data. What is the characteristic property of Poisson distribution? There should be a negligible difference between the mean and variance of the Poisson distribution, if not they are equal. Hence, we try in the next section an incidence rate restricted Poisson (IRRP) distribution which Shanmugam (1991) introduced to the statistics community.

Recall that Y_1, Y_2, Y_3 denote respectively the observed number of confirmed, cured and deaths of coronavirus among their sample spaces. Furthermore, we notice that the random variables (RV) Y_1, Y_2, Y_3 are mutually and significantly correlated with each other from their estimated correlations $\hat{\rho}_{Y_1, Y_2} = 0.96$, $\hat{\rho}_{Y_1, Y_3} = 0.89$ and $\hat{\rho}_{Y_2, Y_3} = 0.79$ are not only positive but also have p-values less than 0.001. The three variables increase or decrease together. The scenario changes when the reality in the field is changing. To realize that let us look at the estimated partial correlation and it is $\hat{\rho}_{Y_2, Y_3 | Y_1} = -0.825$ between the number of cured and the number of deaths by controlling the number of confirmed cases.

It suggests that when the confirmed cases are stable (with no increase or decrease), the number of cured and the number of deaths are inversely proportional. It underlies the importance of flattening the incidences of coronavirus cases. Likewise, let us look at the partial correlation between the confirmed cases and the corona deaths when the number of cured cases become stable (with no increase or decrease). Such partial correlation is estimated to be $\hat{\rho}_{Y_1, Y_3 | Y_2} = 0.921$ and it suggests an undesirable scenario, which indicates that the correlation between the corona deaths will increase when the corona confirmed cases increase and vice versa. In other words, the RV Y_2 , the number of cured cases should not become stable (a flattening level) but rather be increasing in the corona pandemic. These are useful lessons for the government and/or healthcare professionals.

Also, notice in Figure 1 that there is only one spike and it is due to the unusual corona incidences of confirmed as well as deaths in Maharashtra state. Likewise, see in Figure 2 that there are two spikes and it is due to the unusual corona deaths versus cured as they happened in Gujarat and Maharashtra states. Let us proceed further to comprehend the reality in the field. For this purpose, why not build a simple linear regression for predicting Y_3 the number of deaths in terms of the known Y_1 , confirmed and Y_2 , cured cases? Such a regression fit is $\hat{Y}_3 = 34.24 + 0.45y_1 - 6.59y_2$ (with an adjusted coefficient of determination equal to 92%), suggesting that for every 100 new confirmed corona cases, there could be an estimated increase of 45 corona deaths, while for every 100 cured cases, there could an estimated decrease of 659 corona deaths. The importance of curing confirmed corona cases cannot be overstated. Now let us look at the prospects of predictable cure cases in India. For this purpose, we ought to build a simple linear regression for predicting Y_2 the number of cured cases in terms of the known Y_1 , confirmed and Y_3 , corona death cases. The regression fit has happened to be a regression fit is $\hat{Y}_2 = 0.52 + 0.06y_1 - 0.10y_3$ (with an adjusted coefficient of determination equal to 97%), suggesting that for every 100 new confirmed corona cases, there will be only an estimated increase of 6 corona cure cases, while for every 100 corona death cases, there will be an estimated decrease of 10 corona cure cases. The proportionality between the cure and death cases is so frightening to tell the truth and it should not be overlooked by the government or healthcare professionals.

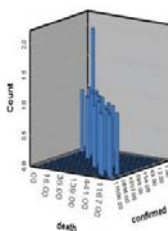


Figure 1. Deaths versus confirmed

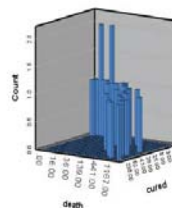


Figure 2/ Deaths versus cured

3. INCIDENCE RATE RESTRICTED POISSON MODEL WITH CORONA CARE EFFICIENCY

Every among the RVs Y_1, Y_2, Y_3 (denoting the number of corona confirmed, cured, and deaths) are nonnegative integers meets the requirements to be a Poisson type with positive *incidence rate* $\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0$ and quantifiable but unknown restriction parameter $\beta_1, \beta_2, \beta_3$ respectively. That is,

$$\Pr[Y_i = y_i | \lambda_i > 0, \beta_i > \lambda_i] = (1 + \frac{y_i}{\beta_i})^{y_i-1} (\lambda_i e^{-\frac{\lambda_i}{\beta_i}})^{y_i} / y_i! e^{\lambda_i}; \quad (1.1)$$

$$y_i = 0, 1, 2, \dots; 0 < \lambda_i \leq \beta_i; i = 1, 2, 3.$$

It has been shown in Shanmugam (1991) that $\Pr[Y_i = y_i | \lambda_i > 0, \beta_i > \lambda_i] \geq 0$ and

$\sum_{y_i=0}^{\infty} \Pr[Y_i = y_i | \lambda_i > 0, \beta_i > \lambda_i] = 1$. Hence, the expression (1.1) is a bonafede probability mass function

(PMF). The PMF (1.1) is recognized as *incidence rate restricted Poisson* (IRRP) distribution. When the restriction is quite liberalized by ineffective social distancing (that is, $\beta_i \rightarrow \infty$), the PMF (1.1) for $i = 1$ transforms to the regular Poisson distribution implying that the outcome y_1 is a rare type. In the regular Poisson distribution, the expected value $E(Y_1) = \lambda_1$ and the variance $Var(Y_1) = \lambda_1$ are equal, which is the characteristic property (see Chattamvelli and Shanmugam, 2019 for details). Otherwise (that is, with an effective social distancing), the number of confirmed corona cases ought to be stable enough, which is echoed in the *stability equation* $E[Y_1 | \lambda_1 > 0, \beta_1 > \lambda_1] = [1 - \frac{\lambda_1}{\beta_1}]^2 Var[Y_1 | \lambda_1 > 0, \beta_1 > \lambda_1]$. The

proportion $[1 - \frac{\lambda_1}{\beta_1}]^2$ is named here a *measure of effective social distancing* index, whose domain is $[0, 1]$.

Once there are a finite nonzero number of coronavirus confirmed cases at a time in a selected location, it makes sense to trace their subsequent status. Though not all of them are hospitalized, several of them who were hospitalized might end up as *cured*, Y_2 . Some of those hospitalized and many of those not hospitalized ends up as *corona death cases*, Y_3 . When there is no restriction on the hospitalization which connects to cure rate (that is, $\beta_2 \rightarrow \infty$), the PMF (1.1) for $i = 2$ transforms to the regular Poisson distribution implying that the outcome y_2 is a rare type. In the regular Poisson distribution, the expected value $E(Y_2) = \lambda_2$ and the variance $Var(Y_2) = \lambda_2$ are equal, which is the characteristic property, (that is, $\beta_2 \rightarrow \infty$). Otherwise (that is, with $\beta_2 \neq \infty$), personifying some restrictions on hospitalization due to a lack of nearby hospitals/clinics, transportation facilities to hospitals etc. in the community), the number of cured corona cases becomes astonishingly less, which is echoed in the *imbalance cure equation*

$E[Y_2 | \lambda_2 > 0, \beta_2 > \lambda_2] = [1 - \frac{\lambda_2}{\beta_2}]^2 Var[Y_2 | \lambda_2 > 0, \beta_2 > \lambda_2]$. The proportion $[1 - \frac{\lambda_2}{\beta_2}]^2$ is named here the *lack of imbalance curability (LIC) index*, whose domain is $[0, 1]$.

In the process, some of those hospitalized and many of that not hospitalized end up as *corona death* cases, Y_3 . When there is no restriction on the corona death rate (that is, $\beta_3 \rightarrow \infty$), the PMF (1.1) for $i = 3$ transforms to the regular Poisson distribution implying that the outcome y_3 is a rare type. In the regular Poisson distribution, the expected value $E(Y_3) = \lambda_3$ and the variance $Var(Y_3) = \lambda_3$ are equal, which is the characteristic property, (that is, $\beta_3 \rightarrow \infty$),. Otherwise (that is, with $\beta_3 \neq \infty$), personifying coronavirus virality level, the person's immunity or resistance level etc.), the number of corona death cases becomes unreliable enough to safety, which is echoed in the *unreliability equation* $E[Y_3 | \lambda_3 > 0, \beta_3 > \lambda_3] = [1 - \frac{\lambda_3}{\beta_3}]^2 Var[Y_3 | \lambda_3 > 0, \beta_3 > \lambda_3]$. The proportion $[1 - \frac{\lambda_3}{\beta_3}]^2$ is named here a *measure of unreliability to safety index*, whose domain is $[0, 1]$.

Following Neyman's $C(\alpha)$ principle (which is a blend of regression and hypothesis testing concepts), Shanmugam (1991) constructed a testing methodology to test the null hypothesis $H_o : \beta = \infty$ against an alternative $H_1 : \beta = \beta_1 < \infty$ and it is based on a statistic, $Z_i = \frac{\bar{y}_i}{s_i^2} \sqrt{(n-1)}$ which is asymptotically normally distributed with mean $\mu_{i,o} = \sqrt{(n-1)} E(\frac{\bar{y}_i}{s_i^2} | H_o)$ and variance $\sigma_{i,o}^2 = (n-1) Var(\frac{\bar{y}_i}{s_i^2} | H_o)$. Hence, the p-value of rejecting the null hypothesis is $p_i - value = \Phi([\frac{\bar{y}_i}{s_i^2} \sqrt{(n-1)} - \mu_{i,o}] / \sigma_{i,o})$. The statistical power of our test statistics is the probability, $1 - \Phi(l)$ of accepting the true alternative hypothesis $H_1 : \beta_i = \beta_{i,1}$ for a specified $\beta_{i,1}$. That is, $l = \{1 - z_\epsilon \sqrt{(1 + \frac{s_y^2}{\bar{y}^3})} / n\} - \mu_{i,1} / \sigma_{i,1}$, where z_ϵ is $100\epsilon^{th}$ percentile of the standard normal distribution.

To find the p-value as well as the power, we need to deal with the moments of the ratio of the correlated sample mean $\bar{y}_i$ and sample variance s_i^2 . Utilizing the formulas to find the expected values and variances of the correlated ratio of random variables in Blumenfeld (2010, pages 8 and 9), we find after algebraic simplifications that

$$\mu_{i,k} = \frac{E(\bar{y}_i | H_k)}{Var(s_{y_i}^2 | H_k)} \left\{ 1 + \frac{Var(s_{y_i}^2 | H_k)}{E(s_{y_i}^2 | H_k)} \right\} - \frac{\rho_{(\bar{y}_i, s_{y_i}^2) | H_k} \sqrt{Var(\bar{y}_i | H_k) Var(s_{y_i}^2 | H_k)}}{\{E(s_{y_i}^2 | H_k)\}^2}$$

and

$$\sigma_{i,k}^2 = \left\{ \frac{E(\bar{y}_i | H_k)}{E(s_{y_i}^2 | H_k)} \right\}^2 \left\{ \frac{Var(\bar{y}_i | H_k)}{[E(\bar{y}_i | H_k)]^2} + \frac{Var(s_{y_i}^2 | H_k)}{[E(s_{y_i}^2 | H_k)]^2} - \frac{2\rho_{(\bar{y}_i, s_{y_i}^2) | H_k} \sqrt{Var(\bar{y}_i | H_k) Var(s_{y_i}^2 | H_k)}}{E(\bar{y}_i | H_k) E(s_{y_i}^2 | H_k)} \right\}$$

where $k = 1, 2$ for the null and alternative hypothesis respectively. Correspond to the null hypothesis,

we substitute $E(\bar{y}_i | H_o) = \lambda_i = Var(y_i | H_o)$, $Var(\bar{y}_i | H_o) = \frac{\lambda_i}{n}$, $Var(s_{y_i}^2 | H_o) \approx \frac{\lambda_i^2}{n}$, and

$$\rho_{(\bar{y}_i, s_{y_i}^2) | H_o} \approx \sqrt{\frac{\{E(\bar{y}_i | H_o)\}^2}{\{E(\bar{y}_i | H_o)\}^2 + \{E(s_{y_i}^2 | H_o)\}^2}} \approx \sqrt{\frac{\lambda_i}{1 + \lambda_i}}$$

value after algebraic simplifications. That is,

$$p_i - \text{value} \approx \Phi \left\{ \left| \frac{\bar{y}_i}{s_{y_i}^2} - 1 \right| \sqrt{n} / \sqrt{1 + \sqrt{\frac{s_{y_i}^2}{\bar{y}_i^3}}} \right\}. \quad (1.2)$$

Likewise, after substituting $E(\bar{y}_i | H_1) = \lambda_i / (1 - \frac{\lambda_i}{\beta_i})$, $E(s_{y_i}^2 | H_1) = Var(y_i | H_1) = \lambda_i / (1 - \frac{\lambda_i}{\beta_i})^3$,

$$Var(\bar{y}_i | H_1) = \lambda_i / n(1 - \frac{\lambda_i}{\beta_i})^3, \quad Var(s_{y_i}^2 | H_1) \approx \lambda_i^2 / n(1 - \frac{\lambda_i}{\beta_i})^6,$$

$$\rho_{(\bar{y}_i, s_{y_i}^2) | H_1} \approx \sqrt{\frac{\{E(\bar{y}_i | H_1)\}^2}{\{E(\bar{y}_i | H_1)\}^2 + \{E(s_{y_i}^2 | H_1)\}^2}} \approx \sqrt{\frac{\{\lambda_i / (1 - \frac{\lambda_i}{\beta_i})\}}{1 + \{\lambda_i / (1 - \frac{\lambda_i}{\beta_i})\}}}$$

and doing algebraic simplifications, we obtain the power function

$$\text{power} = 1 - \Phi \left(\frac{\{1 - z_\varepsilon \sqrt{(1 + \sqrt{\frac{s_{y_i}^2}{\bar{y}_i^3}}) / n - (1 - \sqrt{\frac{\bar{y}_i^3}{s_{y_i}^2 \beta_{i,1}^2})^2} \} \sqrt{n}}{\{(1 - \sqrt{\frac{\bar{y}_i^3}{s_{y_i}^2 \beta_{i,1}^2}})^2 \sqrt{1 + \frac{\sqrt{\frac{s_{y_i}^2}{\bar{y}_i^3}}}{(1 - \sqrt{\frac{\bar{y}_i^3}{s_{y_i}^2 \beta_{i,1}^2}})}\}} \right)} \quad (1.3)$$

Next, we will derive an approximate expression for the survival function for the IRRP distribution, which was not made available in Shanmugam (1991) perhaps because of mathematical intractability. To ease the complexity, we in this article resort Taylorizing only the complexity in the process. That is, the survival function, $S_{y_i}[q, \lambda_i, \beta_i] = \Pr[Y_i \geq q | \lambda_i, \beta_i]$ for the IRRP distribution (1.1) is

$$\begin{aligned}
 S_{Y_i}[q, \lambda_i, \beta_i] &= e^{-\lambda_i} \sum_{y_i=q}^{\infty} (1 + \frac{y_i}{\beta_i})^{y_i-1} (\lambda_i e^{-\frac{\lambda_i}{\beta_i}})^{y_i} / y_i! \\
 &\approx e^{-\lambda_i} \sum_{y_i=q}^{\infty} (1 + \frac{y_i(y_i-1)}{\beta_i})^{y_i-1} (\lambda_i e^{-\frac{\lambda_i}{\beta_i}})^{y_i} / y_i! \\
 &\approx e^{-\lambda_i [e^{-\frac{\lambda_i}{\beta_i}} - 1]} \{ \Pr(\chi_{2qdf}^2 < 2\lambda_i e^{-\frac{\lambda_i}{\beta_i}}) + \frac{(\lambda_i e^{-\frac{\lambda_i}{\beta_i}})^2}{\beta_i} \Pr(\chi_{2[q-2]df}^2 < 2\lambda_i e^{-\frac{\lambda_i}{\beta_i}}) \}
 \end{aligned}$$

due to expanding $(1 + \frac{y_i}{\beta_i})^{y_i-1}$ in terms of $\phi_i = \frac{1}{\beta_i} = 0$ up to the first order term and then reversing to the

parameter β_i and the relationship (see Stuart and Ord, 2015) between the chi-squared distribution function and the cumulative Poisson probabilities indicated. Realize that the chi-squared distribution function is zero when its degrees of freedom (df) is negative. With $q = 1$, an interesting particular result

$$S_{Y_i}[1, \lambda_i, \beta_i] \approx e^{-\lambda_i [e^{-\frac{\lambda_i}{\beta_i}} - 1]} \Pr(\chi_{2df}^2 < 2\lambda_i e^{-\frac{\lambda_i}{\beta_i}}) \quad (1.4)$$

describes the probability of noticing one or more counts in the i^{th} category, where $i = 1, 2, 3$ corresponding to coronavirus *confirmed*, *cured*, or *death* respectively.

The result in (1.4) opens an opportunity to discuss the *odds* for a *jump rate* in the i^{th} category. See Shanmugam and Radhakrishnan (2011) for the importance of the jump rate in the healthcare studies. First, what is called *odds*? The odds of having no episode in the i^{th} category is the ratio $odds_{i, \lambda_i, \beta_i} = \Pr(Y_i = 0 | \lambda_i, \beta_i) / \Pr(Y_i \geq 1 | \lambda_i, \beta_i)$, where λ_i and β_i describe respectively the incidence rate and the impact of the operational restrictions imposed in the i^{th} category. Note that the probability of no episode is $\Pr[Y_i = 0 | \lambda_i > 0, \beta_i > \lambda_i] = e^{-\lambda_i}$, where λ_i is the incidence rate in the i^{th} category.

Consequently, the odds in the i^{th} category, due to (1.11), is $odds_{i, \lambda_i, \beta_i} = \frac{e^{-\lambda_i e^{-\frac{\lambda_i}{\beta_i}}}}{\Pr(\chi_{2df}^2 < 2\lambda_i e^{-\frac{\lambda_i}{\beta_i}})}$ in the

presence of the administrative restrictions. In the absence of restrictions (that is, $\beta_i = \infty$) on the incidence rate in the i^{th} category, the odds become $odds_{i, \lambda_i, \beta_i = \infty} = \frac{e^{-\lambda_i}}{\Pr(\chi_{2df}^2 < 2\lambda_i)}$. Hence, the *odds ratio*

(OR) under the presence versus absence of the administrative restrictions in the i^{th} category is

$$OR_{i, \lambda_i, \beta_i} = \frac{odds_{i, \lambda_i, \beta_i}}{odds_{i, \lambda_i, \beta_i = \infty}} = e^{-\lambda_i [e^{-\frac{\lambda_i}{\beta_i}} - 1]} \left\{ \frac{\Pr(\chi_{2df}^2 < 2\lambda_i)}{\Pr(\chi_{2df}^2 < 2\lambda_i e^{-\frac{\lambda_i}{\beta_i}})} \right\} \quad (1.5)$$

Because $0 < e^{-\frac{\lambda_i}{\beta_i}} < 1$, we notice that the factor $e^{-\lambda_i \{e^{-\frac{\lambda_i}{\beta_i}} - 1\}} > e^{\lambda_i} > 1$ and

$\Pr(\chi_{2df}^2 < 2\lambda_i) > \Pr(\chi_{2df}^2 < 2\lambda_i e^{-\frac{\lambda_i}{\beta_i}})$ since $\lambda_i > 0$. Consequently, we obtain *Theorem 1* below,

whose implications is that the odds under restrictions (that is, $\beta_i \neq \infty$) is more than the odds under no restrictions in every category. In the Category 1, the restrictions promote lockdown and social distancing. In the Category 2, the restrictions impose mandatory admission of coronavirus confirmed individuals. In the Category 3, the restrictions amount to the medical professionals' accessibility to effective medicine, treatment facilities and the encouragement for coronavirus positively tested individuals to increase their immunity level and corona resistance health.

Theorem 1. The odds of not getting coronavirus positive test results in Category 1, being not admitted in a hospital in Category 2, or not dying due to coronavirus under restrictions is more than the analogous odds under no restriction.

Furthermore, the discrete hazard rate, $h_i(1|\lambda_i, \beta_i) = 1 - \frac{S_{y_i}(2|\lambda_i, \beta_i)}{S_{y_i}(1|\lambda_i, \beta_i)} = 1 - \frac{\Pr(\chi_{4df}^2 < 2\lambda_i e^{-\frac{\lambda_i}{\beta_i}})}{\Pr(\chi_{2df}^2 < 2\lambda_i e^{-\frac{\lambda_i}{\beta_i}})}$ under

restrictions in comparison to the hazard rate,

$h_i(1|\lambda_i, \beta_i \neq \infty) = 1 - \frac{S_{y_i}(2|\lambda_i, \beta_i \neq \infty)}{S_{y_i}(1|\lambda_i, \beta_i \neq \infty)} = 1 - \frac{\Pr(\chi_{4df}^2 < 2\lambda_i)}{\Pr(\chi_{2df}^2 < 2\lambda_i)}$ under no restrictions. These results are

summarized in *Theorem 2* below. See Shanmugam and Chattamvelli (2015) for definition and usefulness of the hazard rate.

Theorem 2. The reduction, R_i in the hazard rate due to the restriction scenario compared to the no restriction scenario is

$$R_i = \Pr(\chi_{4df}^2 < 2\lambda_i) \{\Pr(\chi_{4df}^2 < 2\lambda_i)\}^{-1} - \Pr(\chi_{4df}^2 < 2\lambda_i e^{-\frac{\lambda_i}{\beta_i}}) \{\Pr(\chi_{2df}^2 < 2\lambda_i e^{-\frac{\lambda_i}{\beta_i}})\}^{-1}. \quad (1.6)$$

Since the variance describes heterogeneity, a measure of vulnerability in the i^{th} category could be

defined as $Vulnerability_i = \frac{Var[Y_i | \lambda_i > 0, \beta_i > \lambda_i]}{Var[Y_i | \lambda_i > 0, \beta_i > \lambda_i] + E[Y_i | \lambda_i > 0, \beta_i > \lambda_i]}$ and it simplifies to

$Vulnerability_i = \{1 + [1 - \frac{\lambda_i}{\lambda_i + \delta_i}]^2\}^{-1}$, where $\delta_i = \beta_i - \lambda_i$ is the gap representing the impact level of

the restrictions to reduce the incidence rate λ_i . It suggests that in the absence of restrictions *that is,

$\delta_i = 0$) in the i^{th} category, the vulnerability is 100%. Otherwise (that is, in the presence of restrictions, as $\delta_i > 0$), reduces the vulnerability. Hence, there is a reason to designate

$$VI_i = \{1 + [1 - \frac{\lambda_i}{\lambda_i + \delta_i}]^2\}^{-1} \quad (1.7)$$

as the vulnerability index in the i^{th} category.

The coronavirus (for that matter, any pandemic) is undesirable and is to be averted. Is the pandemic at least stable enough for the healthcare professionals, administrative agencies, and most importantly the public feel quite comfortable to deal with it? If so, how do we define and measure the stability level of the pandemic? One way to define its *instability level* is to utilize the survival function (1.4). In other words, if the uncertain chance mechanism behind the corona virus hold a memory, it is likely to exhibit it terms of an instability. We define the *stability index* (SI_i) for the i^{th} category as

$$\begin{aligned} SI_{i,\beta_i} &= e^{-\lambda_i [e^{-\{\frac{\lambda_i}{\beta_i}\}} - 1]} [\Pr(\chi_{4df}^2 < 2\lambda_i e^{-\{\frac{\lambda_i}{\beta_i}\}}) - e^{-\lambda_i [e^{-\{\frac{\lambda_i}{\beta_i}\}} - 1]} \{\Pr(\chi_{2df}^2 < 2\lambda_i e^{-\{\frac{\lambda_i}{\beta_i}\}})\}^2] \\ &\approx e^{-\lambda_i [e^{-\{\frac{\lambda_i}{\beta_i}\}} - 1]} \Pr(\chi_{4df}^2 < 2\lambda_i e^{-\{\frac{\lambda_i}{\beta_i}\}}) \end{aligned} \quad (1.8)$$

Under restrictions in comparison to what it is $SI_{i,\beta_i=\infty} = \Pr(\chi_{4df}^2 < 2\lambda_i) - \{\Pr(\chi_{2df}^2 < 2\lambda_i)\}^2$ under no restriction. Hence, the improvement due to imposing restrictions in the i^{th} category is $SI_{i,\beta_i} - SI_{i,\beta_i=\infty}$.

The *Tail-Value at Risk* ($TVaR_i$) is a commonly used concept in the business world (see Khokhlov, 2016 for details). and it can be fitting here to describe the risk at an outbreak of pandemic like coronavirus. It is

$$\begin{aligned} TVaR_{\beta_i,1} &= E[Y_i | Y_i \geq 1, \lambda_i > 0, \beta_i > \lambda_i] = 1 + \frac{\sum_{y_i=1}^{\infty} (y_i - 1) \Pr[Y_i \geq 1 | \lambda_i > 0, \beta_i > \lambda_i]}{\Pr[Y_i \geq 1 | \lambda_i > 0, \beta_i > \lambda_i]} \\ &= \frac{e^{-\lambda_i [e^{-\{\frac{\lambda_i}{\beta_i}\}} - 1]}}{\Pr(\chi_{2df}^2 < 2\lambda_i e^{-\{\frac{\lambda_i}{\beta_i}\}})} \end{aligned}$$

under restrictions in comparison to $TVaR_{\beta_i=\infty,1} = \{\Pr(\chi_{2df}^2 < 2\lambda_i e^{-\{\frac{\lambda_i}{\beta_i}\}})\}^{-1}$ under no restriction for i^{th} category. In other words, we obtain a theorem below.

Theorem 3. The value at risk, $TVaR_{\beta_i,1}$ under restrictions is reduced by a percent $e^{-\lambda_i [e^{-\{\frac{\lambda_i}{\beta_i}\}} - 1]}$ from that of the value at risk, $TVaR_{\beta_i=\infty,1}$ at the outbreak of coronavirus.

The estimable reduction in the reduction of tail-value at risk in the i^{th} category is therefore

$$\hat{\mathfrak{R}}_i = e^{-\lambda_i [e^{-\{\frac{\lambda_i}{\beta_i}\}} - 1]} = e^{-\sqrt{\frac{\hat{y}_i^2}{s_y^2}} [e^{-\{1 - \sqrt{\frac{\hat{y}_i^2}{s_y^2}}\}} - 1]} \quad (1.9)$$

In the next section, we illustrate all the derived expressions so far using the coronavirus data in the open domain webpage <http://mohfw.gov.in>.

4. ILLUSTRATION WITH CORONA CASES IN INDIA

In this section, we illustrate the expressions derived with India's coronavirus data which are displayed in Table 1. The data denote the number, y_1 of *confirmed* cases, the number, y_2 of *cured* cases and the number, y_3 of *death* cases in $n = 32$ states/territories as of May 1, 2020. Of course, the counts change on daily basis. Their average, $\bar{y}_i$ and variance, $s_{y_i}^2$ are calculated and displayed in Table 1. Notice that the sample average is less than the sample variance in each category, which is the characteristic property of the *incidence rate restricted Poisson* (IRRP) distribution (1.1). The maximum likelihood estimate (MLE) of the *incidence rate* and the *restriction* parameter are $\hat{\lambda}_i = \sqrt{\frac{\bar{y}_i^3}{s_{y_i}^2}}$ and $\hat{\beta}_i = \frac{\hat{\lambda}_i}{(1 - \sqrt{\frac{\bar{y}_i}{s_{y_i}^2}})}$

respectively. The estimates are displayed in Table 1. There is a finite amount of effect due to restrictions, according the above parameter estimates. The MLE, $\hat{\delta}_i$ of the *gap* using the expression $\hat{\delta}_i = \hat{\beta}_i - \hat{\lambda}_i$ are then computed and displayed in the Table 1. They are $\hat{\delta}_1 = 0.27$, $\hat{\delta}_2 = 0.46$, and $\hat{\delta}_3 = 0.17$ in the *confirmed*, *cured* and *death* category respectively. According to (1.2), the p-value for the null hypothesis $H_o : \beta = \infty$ to be true is calculated and displayed for each category in Table 1. The null hypothesis is less likely to be true in all categories, implying that there is a finite and significant effect due to the restrictions imposed in each category. Recall that the restriction level, β_i is estimated to be 18.20, 12.34, and 2.66 respectively in the corona *confirmation*, *cure*, and *death* category. In an event of achieving the restriction level $\beta_{i,1} = 25.00$, how likely our methodology will accept it with the data in Table 1 is the *statistical power* and it is calculated and displayed in Table using expression (1.3). The power is 0.37, 0.12, and 0.99 in the coronavirus *confirmed*, *cured* and *death* category, respectively. The *odds ratio* in (1.5) is one across all categories, meaning the situation is quite even under the presence to the absence of the administrative restrictions. According to the *Theorem 2*, and (1.6), the reduction, R_i in the *hazard rate* due to the restriction scenario compared to the no restriction scenario is 7.69, 5.53, and 1.98 in the coronavirus *confirmed*, *cured* and *death* category, respectively, impressive enough to feel that the hazard reduction is significant in the death category. The measure of *vulnerability* in (1.7) is calculated and displayed in Table 1. The vulnerability measure is 0.99, near its maximum in all the categories warning us the seriousness of coronavirus virality. The stability index in (1.8) is tiny

alerting that the scenario of coronavirus spreading is volatile. The reduction (1.9) in the tail-value at risk in the th category is extremely high, medium, and comfortably less in the corona *confirmation*, *cure*, and *death* category respectively, suggesting that the curing as well as confirming categories demand much more improvements during the coronavirus pandemics.

5. CONCLUSION

To devise more strategies to reduce (if not eliminating totally and ideally) the ugly impacts of the coronavirus in community, much better understanding of the existing and destined ideal situation is currently a necessity. For such a comprehensive and better understanding, the coronavirus database should be expanded to have details on related covariates. Also, more regression type suitable methodologies need to be developed, demonstrated, and recommended to the advising groups to the decision makers like the local, state, national, and global administrative agencies.

6. REFERENCES

- Blumenfeld, D., 2010, *Operations Research Calculations Handbook*, CRC Press, Boca Raton, Florida.
- Chattamvelli, R. and Shanmugam, R., 2019, *Generating Functions in Engineering and the Applied Sciences*, Synthesis Lectures on Engineering, Morgan & Claypool Publishers, 82 Winter sport Lane, Williston, VT 05495, USA.
- Shanmugam, R., 1991, Incidence rate restricted Poissonness. *Sankhya (B)*, 53, 2, 191-201.
- Shanmugam, R. and Chattamvelli, R., 2015, *Statistics for Scientists and Engineers*, John Wiley Inter-Science Publication, Hoboken, New Jersey 07030, USA.
- Shanmugam, R. and Radhakrishnan, R. (2011), Incidence jump rate reveals over/under dispersion in count data, *International Journal of Data Analysis and Information Systems*, 3, 1, 1-8.
- Stuart, A. and Ord, K., 2015. *Kendall's Advanced Theory of Statistics, Volume 1*. Oxford University Press, London.