Appendix

Technical Note

1. Model based simulation methods for count data: Estimation of counterfeits involves modelling of count data observed for a moving population where nobody can put a conceivable upper bound for a given space and time domain. It is akin to tracking occurrence of a particular trait-related event (contingent event) in a sequence of experimental trials. Examples are accidents on a highway, number of misprints or defective items, number of birds, fishes etc. of specific species in the population based on a series of hauls. Model based simulation of count data finds intensive applications in the areas of clinical trials, estimation of abnormality of blood cells in haematology, cytogenetic studies on chromosomal abnormalities and enumeration of aquatic marine species of rare variety. Crash count data analysis also is an important area being pursued in traffic safety analyses. In such phenomena, population sizes could be large as well as variable, but mostly unknown where the entities can be sighted only at a very low frequency. The first statistical method of estimation for such low frequency event for large population was made by J.B.S. Haldane (1945) based on sequential sampling. Two specific features of such population are notable for modelling purpose. One is low but varying nature of frequency of occurrence of the particular trait-related event and the other aspect is that the variance of the expected number of occurrence of the event is more than the expected number. Inverse sampling is an often-used method adopted for estimating frequency of occurrence of such low probability event in a highly dispersed population.

1.1. Classical approach (Fixed sample size**)**: When frequency (π) of the key attribute does not change much from one sample to another, *fixed sample size* approach is suggested to estimate π, for which the standard error estimate (SEE) of the sample estimate (p = number of sampling units bearing the attribute/n) is $\sqrt{(\pi (1-\pi)/n)}$. It is akin to tossing a coin *n* times (Bernoulli trials) and observing number of 'head's (generally termed as 'success'). Each trial (or, tossing the coin) is assumed to be carried out independently where chance of getting a 'head' is π and the resulting distribution model is *Binomial Distribution*. Such trial runs of prefixed finite sizes would however lead to biased estimate if it is felt that π is varying in nature, particularly when occur with some degree of rarity. If n = 1000, and π = 0.3, the SEE is 0.015, but if π = 0.01, the SEE is 0.0031. Very low SEE makes any two different populations indistinguishable with more and more smaller π. For example, when π =0.01 and 0.005, SEE= 0.0031 and 0.0022 respectively. Similarly, for all the more

smaller frequencies π = 0.000573, 0.000533, 0.000491, 0.000327 and .000263, if taken as tentative frequencies of counterfeits as commonly reported ratio of counterfeit notes observed amidst a very large size of currency in circulation), the SEE = 0.00076, 0.00073, 0.00070, 0.00057 and 0.00051 respectively, which are very small rendering the estimation procedure statistically meaningless.

1.2. Inverse sampling (Variable sample size**)**: Besides low value of π, very large population size and possible non-stationary nature of variability of π over time would render the fixed sample size procedure to estimate frequency of count data untenable. Sudden spurt in counterfeiting activities leading to sizable jump in counterfeit notes detection may alter the occurrence of the contingency of encountering a forged note and in such situations inherent variability in the count data could be more than expected number of trials required to detect any fixed number of counterfeits. To have a meaningful estimation procedure, J.B.S Haldane (1945) had introduced the method of inverse (binomial) sampling based on probability distribution based method which requires that random sampling be continued until a specified quota of units with the attribute (counterfeit) has been obtained. The method is based on distributional model that suits the empirically observed property of relevant count data. If the proportion of individuals possessing a certain characteristic is π and we sample until we see *r* such individuals, then the number of individuals sampled is a negative binomial random variable. Its relevance can be further understood from the following three alternative models used for infinitely large count data.

1.3. Simulation of large binomial count data: Inverse sampling method is premised upon three commonly adopted frequency/density estimation models for infinitely large binomial count data namely (i) *Poisson Distribution*, (ii) *Negative Binomial Distribution* and (iii) *Negative Hypergeometric Distribution*.

1.3.1 Poisson probability models: Single parameter (λ) *Poisson model* is the distribution of the number X of certain random events occurring in the course of a sequence of trials where frequency function is $P\{X = k\} = e^{-\lambda}(\lambda)^{k}/k!$, k = 0, 1, 2 …. When used for modelling the distribution of random number of points occurring in a pre-designed area, the parameter λ of the distribution is proportional to the size (length, area or volume) of the domain. Then, λ is the expected number of the contingent event (rate per unit of time, say a month or a year) and k is the sample observations on number of discrete the events recorded in the experimental trial. Poisson distribution gives a fair approximation to binomial distribution connected with

a sequence of fixed number of independent trials yielding to 'success' or 'failure' in each trial (e.g., 'head' or 'tail' in tossing a coin *n* times; 'success' may be termed as sighting/detecting 'counterfeit' while inspecting a *pre-fixed* number (*n*) of currency notes). Poisson probability law works as a good approximation when n is large and very small chance (π) of occurrence of 'success' (detecting a 'counterfeit') so that *np* is more or less a stable number ($\approx \lambda$). Adopting Poisson models has an overriding requirement, namely Mean = Variance ($=\lambda$). However, in reality it is often found that variance is larger than the average value observed empirically, which actually characterises a dispersed population. Then the commonly adopted approach is to fit a negative binomial distribution. It may be noted, even in finite case of binomial distribution, the expected value is larger than the variance ($n.\ \pi > n.\ p.\ (1-\pi)$). Neither fixed large sample sized Binomial distribution nor Poisson distribution (as a limiting form of large but fixed size Binomial distribution) is suitable in such cases. Negative binomial fits the situation as theoretically the variable of interest (notes in circulation) can be infinitely large and the variance is greater than the expected value of the variable.

1.3.2. Modelling over-dispersion: Over-dispersion is a typical feature encountered in large size count data that is not amenable to Poisson distribution based model simulation. Ignoring dispersion amounts to overweighing the data and consequent underestimating the uncertainty. The well-known Poisson distribution is fully definable by a single parameter, the mean (λ), which is equal to its variance. But as would be discussed below, variance to mean ratio could significantly exceed unity, which is often referred to as over-dispersion. Many such count data are satisfactorily fitted with the negative binomial distribution (NBD), which finds ready applications to various biological and industrial problems. Student (1907) derived its distribution during the course of making counts of yeast cells. Subsequently scores of applied researchers established successfully various forms of the negative binomial model in explaining counts of insects pests, problems of germination records in the 1940s and 1950s or even better modelling of dispersion parameter being pursued in the recent time for motor collision data with low sample mean obtained for small sample sized observations.

Here lies the ingenuity of designing the experiment of observing the data based on occurrence of the particular entity or event being tracked through (i) suitable structured area or zone (e.g., dividing sea-bed in square units to observe presence of an aquatic specimen, which we call a 'success' amidst remaining other species termed as 'failure'), or (ii) different time periods as well as zones (e.g., selected peak hour periods for important part of vehicular traffic lanes and crossings to observe number of accidents, the so called 'successes', against vehicles passing through

without any accident). There could be bunching of events i.e., occurrence of multiple 'successes', in real life phenomenon because of scaling problem, particularly if it happens to be oft repeated a phenomenon. Even with a workable scaling, modeling efforts may need to data censoring technique or identifying mixtures of probable underlying random behavior. All these are very much true foe estimating counterfeit currency notes circulating in the system.

1.3.3. Negative binomial distribution (NBD): NBD is used for simulating count data pertaining to occurrence of dichotomous outcomes in the form of 'success' or 'failure' depending on sighting of the particular trait successfully or not in a sequence of independent trials. By assigning the probability of success (say, π) in each trial, if experiment continues until a total of *r* successes are observed, where *r* is fixed in advance, the probability distribution of the number of failures (**X**) before the ***r***-th success follows the following frequency law[21]:

$$P(X = x) = \binom{x + r - 1}{r - 1} \pi^r (1 - \pi)^x, \qquad (10)$$

Here, $0 \leq \boldsymbol{\pi} \leq 1$, **x** = 0, 1, 2…., and the random variable X denotes number of failures before the ***r***-th success is observed. For example **X** = 0 means all the first ***r*** trials have resulted in a continuous chain of r successes; and for observation like **X** = k, it means that n = k + ***r*** number of trials have led to r number of successes. We would denote the fact that '**X** is distributed as negative binomial distribution with the parameters r and π' as **X** ~ NB (***r,*** π).

Some useful properties of the negative binomial distribution are worth mentioning here. (i) Mean, variance and skewness of a distribution are critical for modeling count data on occurrences of counterfeit notes. Mean (i.e., expected value in the form of arithmetic mean or common average term) and the variance of NB (***r,*** π) are:

$$E(X) = \frac{r(1 - \pi)}{\pi}, \quad \text{Var}(X) = \frac{r(1 - \pi)}{\pi^2}.$$

The distribution is positively skewed meaning that the right tail is longer with the *mass* of the distribution getting concentrated on the left of the figure. It has a few relatively high values so that mean > median > mode. As per standard measure of skewness, negative binomial distribution portrays very high amount of positive skewness: $(2 - \pi)/ \sqrt{(\boldsymbol{r}\ (1 - \pi))}$. It reduces with increased ***r*** and becomes almost symmetric for large ***r*** (≥ 40) like the bell-shaped normal curve whereby mean ~

[21] The formula for P(X = x) in (10) can be written as $(-\mathbf{1})^x \binom{-r}{x}\ \pi^r\ (\mathbf{1} - \boldsymbol{\pi})^x$, the r-th term of $\sum_{x=0}^{\infty} \binom{-r}{x} \pi^r\ (\mathbf{1} - \boldsymbol{\pi})^x = \pi^r\ (\mathbf{1} - (\mathbf{1} - \boldsymbol{\pi}))^{-r}\ = \mathbf{1}$, which involves 'negative binomial' terms.

median ~ mode and for moderately large ***r***, it behaves like a Poisson distribution with the mean rate (λ) ~ ***r*** $(1 - \pi)/ \pi$. The parameter r, known as the *shape parameter*, helps model the underlying distribution flexibly so as to include variety of possible shapes within generally acceptable ranges of moderately small values of ***r***. Key terms and properties associated with NBD-based simulation[22] are as under.

1.3.3.1 Measure of dispersion: Variance of the NBD random variable is greater than its expected value, which is a key feature of the distribution. As Var (X)/E (X) = $1/\pi$ (>1), departure of variance to mean from 1 makes the occurrence of successes more sparse or dispersed for lower chance of occurrence of success. In reality, to make the cont data amenable to NBD model, the sample data ought to exhibit a large variance and a small mean, and display over-dispersion with a variance-to-mean value greater than one.

- **Dispersion parameter**: φ = Var (X)/E(X) – 1 = $(1/\pi - 1) = (1-\pi)/ \pi$ gives a measure of dispersion of relatively rare trait in the count data. (It is mostly interpreted as the degree of departure from orderly behaving Poisson distribution).
- **NBD based inverse sampling**: To adopt *inverse sampling* scheme by fitting NBD to observed frequency data on 'successes' namely, say, detecting counterfeit notes in a sequential random draw of currency notes in circulation (NIC), dispersion (φ) of the count data has to be large. As real life simulation exercises are concerned, for very small value of π, the chance of occurrence of one unit of counterfeit, one may need to inspect a very large number of notes in circulation (NIC).
- **The term "inverse"**: If $X_r \sim$ NB D(***r,*** π) and for any fixed s, Y_{s+r} is the random variable representing the *binomial distribution* with parameters *s* + *r* and π, then it can be shown that: Pr ($X_r \leq s$) = Pr ($Y_{s+r} \geq r$) → Probability that there are at least r successes out of s + r trials. (It may be noted that X_r can take very large integral value, whereas Y_{s+r} is of finite size from 0 to s + r). In this sense, the negative binomial distribution is the "inverse" of the binomial distribution.
- **Estimation**: Suppose π is unknown and an experiment is conducted where it is decided ahead of time that sampling will continue until *r* successes are found. The sufficient statistics for the experiment is the number of failures (*k*).

[22] Ideally speaking, sampling without replacement case fits the counterfeit note examination case, for which negative hypergeometric distribution based simulation would be an ideal approach. However, for relatively large numbers, it is better to approximate with negative binomial distribution. Computing variance and higher order moments is an involved exercise which requires much compuatation intensive process which can be adopted for further fine-tuning.

In estimating π, the minimum variance unbiased estimator (MVUE) derived by Haldane (1945) is: $\hat{p} = \frac{r-1}{r+k-1}$, and not the common sense estimator, namely $\tilde{p} = \frac{r}{r+k}$, because this is biased.

- **Best estimator:** Minimum variance unbiased estimator (MVUE) of frequency of occurrence of the less common attribute ('success) is, therefore, $\hat{p} = \frac{(r-1)}{(n-1)}$ and *not r/n*, the ratio of the number of 'successes' to the total number of trials where n = r + k, as is the case for a fixed and finite number of Bernoulli trials. This $\hat{p}$ is also minimum variance unbiased estimator (MVUE). An unbiased estimator of the variance of the above MVUE $\hat{p}$ was shown by Finney (1949) to be $\hat{p}\,(1-\hat{p})/(n-2)$. Therefore, in **inverse sampling** simulation based on negative binomial distribution (NBD), the MVUE of the success probability and unbiased estimator of its variance are:

$$\hat{p} = \frac{(r-1)}{(n-1)}, \qquad \text{Estimate of Varince}\,(\hat{p}) = \frac{\hat{p}(1-\hat{p})}{(n-2)}$$

- **Practical rules**: Following observations about the MVUE $\hat{p}$ and estimator of its variance are useful for having some practical rules for adopting inverse sampling.
  - The standard error is a satisfactory indicator of the error of estimation of π only when r is large.
  - Actually, Var ($\hat{p}$) has a complicated expression and sharper bounds have been reported in the literature and some subsequent works[23]. Assuming *s* = Var ($\hat{p}$), **an upper limit to s/$\hat{p}$ can be fixed in advance of sampling for a reasonable value of r**, so that an upper value can be obtained for this ratio as s/$\hat{p}$ = √ {(1-$\hat{p}$)/(r – 1 –$\hat{p}$)} ≈ √ {1/(r-1)} for small $\hat{p}$.
- **Randomness and error estimate**: As the practice of inverse sampling involves collection of the data in order, the sample also provides evidence of whether the condition of independence of successive observations is fulfilled. If successive individuals are independent of one another, (r-1) entities having the attribute should be distributed *at random* intervals throughout the first (n-1) counted; a departure from independence, such as would result from a clustering of counterfeits) would increase the frequency of short and long gaps between these intervals at the expense of intervals of moderate length. A test of significance may

[23] The paper on "Estimation of a probability with optimum guaranteed confidence in inverse binomial sampling" by Luis Mendo and Jos'em. Hernando, Bernoulli 16 (2), 2010, 493-513 provides the latest update on the matter.

be carried out periodically based upon the observed frequency with which the units with the 'rare trait' are preceded and followed by the normal ones or some other suitable statistic based upon the length of intervals. Significant deviation from the value predicted by a hypothesis of randomness of intervals, whether resulting from clustering of the less common phenomenon from an expected regularity of intervals, would indicate that the standard error and limits of error cited above may not be applicable (Finney, 1949).

- **Stylized properties of NBD**
    - Additive property: If finitely many X_i's are independently distributed as NBD (r_i , π), then $\sum X_i$ is distributed as NBD ($\sum r_i$, π).
    - Limiting property: The negative binomial distribution is better approximated by the Poisson Distribution in the following sense:

$$\text{Poisson}(\lambda) = \lim_{r\to\infty} \text{NegBin}(r, r/(\lambda + r)).$$, where

$$\lambda = r.(\pi^{-1} - 1) \text{ and } \pi = r/(r + \lambda).$$

    - If X_r is a random variable following the negative binomial distribution with parameters *r* and π, then X_r is a sum of *r* independent variables following the NBD (1, π) with parameter π. As a result, X_r can be approximated by *normal distribution* for sufficiently large *r*. Dimensionally r = 40 to 50 may treated as large to use an overall Poisson approximation to explain binomial count data reasonably well. Beyond that, normal distribution could be invoked as a limiting case. In practice, observing incidence of counterfeiting may be enough to restrict below 20. Practical range found to be ranged between 4 to5, which needs to be firmed up based on empirical exercises. In case of real life problems practitioners adopt *re-parametrised* version of the classical $X \sim$ NB D(r, π) model, which is expressed in terms of mean (m = r φ) and the shape parameter (r), or equivalently in terms of its mean and variance as may be denoted as RNBD (m, m + m^2/r). In such representation, r is termed as shape parameter and φ is called dispersion parameter.

1.3.3.2 Re-parametrised versions of NBD: Following transformed versions of the above classical form of Negative Binomial distribution add to the interpretation power and model explanation when fitted to empirical data. The most common transformed versions are as under.

- Conventionally a transformed version of negative binomial distribution is referred as Pascal distribution: If $X \sim$ NB D(***r,*** π) with X = 0, 1, 2,…., then Y =

X + r is termed as Pascal distribution, which denotes number of independent trials required to observe r 'successes'. Its probability mass function is

$$P(Y = n) = \binom{n-1}{r-1} \pi^{r} (1 - \pi)^{n-r}, \quad \text{where } n = r, r+1, r+2, \ldots\ldots$$

- **Commonly used transformed version** of inverse or negative binomial distribution by the practitioners is in terms of two parameters namely mean (m = r φ) and the shape parameter (r). The original form of the NBD takes the following transformed version:

$P(X = x) = \binom{x+r-1}{r-1} (m/(m+r))^{x} (1+m/r)^{-r}$, $x = 0, 1, 2, \ldots\ldots\ldots, m>0, r>0$. This version[24] is denoted as NBD (u, r) in terms of mean and shape parameter (r), which helps interpret the distribution in terms of dispersion (φ = u/r) and shape parameter (r) directly. When expressed in terms of mean (u) and the shape parameter (r), the mean and variance are: $m = r. (1 - \pi)/\pi$ and variance = $m + m^2/r$. So the dispersion parameter = $m/r = (1/\pi - 1)$. For fixed m (average number of failures before r-1 successes), dispersion and shape parameter is inversely related. Though some practitioners termed the reciprocal of the shape parameter (1/r) as dispersion parameter as commonly reflected from an average recurrent pattern in the randomly occurring sequences of 'success' and 'failure', dispersion is better represented by the underlying low chance (π) of occurrence of a 'success'.

1.4. Examples: (i) In case of a rifle range with an old gun that misfires 5 out of 6 times, if one defines ``success'' as the event the gun fires; if X is the number of failures before the third success, X ~ NBD (3**,** 1/6). The probability that there are 10 failures before the third success is given by

$$P(X = 10) = \binom{12}{2} \left(\frac{1}{6}\right)^{3} \left(\frac{5}{6}\right)^{10} = 0.0493.$$

The expected value and variance of *X* are

$$E(X) = \frac{r(1-\pi)}{\pi} = \frac{3 \cdot \frac{5}{6}}{\frac{1}{6}} = 15,$$
$$\mathrm{Var}(X) = \frac{r(1-\pi)}{\pi^2} = \frac{3 \cdot \frac{5}{6}}{\frac{1}{36}} = 90.$$

As mentioned above, finite size Binomial distribution modeling and its large size but moderate π as approximated by Poisson distribution does not suit the occasion

[24] This alternative stylized form is used in current literature where variance is m (1 +m/r), where decreasing values of 'r' correspond to increasing dispersion for stable (fixed) value of 'm' (ref. Lloyd-Smith, James O (2007)).

where it may be visibly evidenced that sample estimate of variance exceeds the sample mean. This leads to the problem of estimation in case of over-dispersed count data, which can be best understood from some real life example.

(ii) A real life example: To understand the incidence of dispersion problem, let us cite an empirical example on observing aquatic invertebrates from selected samples of marine lives. The Table below gives the number of aquatic invertebrates on the bottom in 400 square units. (It is not time indexed but is an area based sampling).

No. of aquatic invertebrates (x)	**0**	**1**	**2**	**3**	**4**	**5**	**Total**
Frequency (f)	213	128	37	18	3	1	400

The above Table is a typical presentation adopted for enumerating count data for tracking event like observing specific marine species per square units. Here, the experiment comprises dividing any select 400 square units of sea bed into 400 squares, each of unit area size and observing number of invertebrates per square units. So f = 213 means so many unit-sized squares are observed to be without any invertebrates. Similar example of observing counterfeits over select time period (say 52X4 =208 serially arranged weekly indexed four zones of the country) can be constructed. The fitting exercise with negative binomial distribution for the transformed version in terms of diffusion and shape parameters are described below.

Estimated mean and variances are: $\bar{\mathbf{x}} = \bar{\mathbf{m}} = \frac{1}{\sum_i \mathbf{fi}} \sum_i \mathbf{fixi} = \frac{273}{400}$ **= 0.68** and

$$s^2 = \tilde{\sigma}^2 = \frac{1}{\sum_i f_i - 1}\left[\sum_i f_i x_i^2 - \frac{(\sum_i f_i x_i)^2}{\sum_i fi}\right] = \frac{1}{399}\left[511 - \frac{273^2}{400}\right] = \mathbf{0.81}$$

We will use this empirical data to fit NBD (m, r). The empirical exercise cited below is based on some alternative notations. Let $\tilde{q}\ (= \frac{\tilde{\sigma}^2}{\bar{m}})$ be an estimate of 1/π, $\tilde{q} = \tilde{p} + 1$ 1 where $\tilde{p}$ is an estimate of the dispersion parameter (φ). Then $\tilde{\sigma}^2 = \tilde{r}.\tilde{p}.\tilde{q} = \tilde{m}.\tilde{q}$ and we get 0.81 = 0.68.$\tilde{q}$; $\tilde{q} = 1.19, \quad \tilde{p} = 0.19.$

$\tilde{k} = \frac{\tilde{m}}{\tilde{p}} = \frac{0.68}{0.19}$ = 3.58 and the estimated probabilities (relative frequencies**)** are:

$$P(x=0) = q^{-r} \text{ and } P(x+1) = \left(\frac{\tilde{r}+x}{x+1}\right)\left(\frac{\tilde{m}}{\tilde{m}+\tilde{r}}\right) P(x).$$

[(i.e, **P(x=0)** = $(1.19)^{-3.58}$ = 0.5365; **P(x=1)** = (3.58/1) X (0.1596) X (0.5365) = 0.3065; **P(x=2)** = ((3.58+1)/2)X(0.1596)X(0.3065) = 0.1120; **P(x=3)** = ((3.58+2)/3)X(0.1596) X (0.1120) = 0.0332; **P(x=4)** = ((3.58+3)/4)X(0.1596)X(0.0332)=0.0087; **P(x=5)** = ((3.58+4)/5)X(0.1596)X(0.0087)=0.0022 X $\sum_{x=0}^{5} P(x)$ = 1).

Estimated theoretical frequencies (N.B.D.): $N_{x=0}$ = 400 × P(x=0) = 400 × 0.5365 = 214 $N_{x=1}$ = 400 × P(x=1) = 400 × 0.3065 = 123, $N_{x=2}$ = 400 × P(x=2) = 400 × 0.1120 = 45, $N_{x=3}$ = 400 × P(x=3) = 400 × 0.0332 = 13, $N_{x=4}$ = 400 × P(x=4) = 400 × 0.0087 = 4. $N_{x=5}$ = 400 × P(x=5) = 400 × 0.0022 = 1, $\sum_{x=0}^{5} N_x = 400$].

Testing goodness of fit: A problem that arises frequently in statistical work is the testing of comparability of a set of observed (empirical) and theoretical (N.B.D.) frequencies. To test the hypothesis of goodness of fit of the NBD to the empirical frequency distribution we calculate the

value of $\chi^2 = \sum \frac{(fi - \theta i)^2}{\theta i}$, where f_i = empirical frequencies and θi = theoretical frequencies. The estimated χ^2 - value is compared with the tabulated $\chi^2_{u,\,a}$ -value. The hypothesis is valid if $X^2 < \chi^2_{u,\,a}$, the hypothesis is discredited if $X^2 > \chi^2_{u,\,a}$.

(N.B: It should be noted that, since χ^2 curve is an approximation to the discrete frequency function; care must be exercised that the χ^2 test is used only when the approximation is good. Experience and theoretical investigations would justify whether the approximation is satisfactory or not. The following Table gives the empirical and theoretical frequencies of the previous example and the estimated χ^2 value.)

Table: χ^2 - test of goodness of fit negative binomial distribution (NBD) to spatial distribution of aquatic invertebrates

Number of aquatic invertebrates (x)	**Number of squares**		**$(f_i - q_i)$**	$\frac{(f_i + \theta_i)}{\theta_i}$
	Empirical frequencies (f_i)	Theoretical frequencies (q_i)		
0	213	214	-1	0.0047
1	128	123	+5	0.2033
2	37	45	-8	1.4222
3	18	13	+5	1.9231
4	4	5	-1	0.2000
				X^2 =3.7533

The tabulated value of $\chi^2_{\nu=2,\,a=0.05}$ = 5.991

(ν - degrees of freedom, ν = 5 classes - (2 estimated parameters + 1))

Since $X^2 < \chi^2_{u,a}$ (i.e.,3.7533 < 5.991), the hypothesis of goodness of fit is valid.

(N.B: A second estimate of (r) can be obtained as $\tilde{\sigma}^2 = m + \frac{m^2}{r}$, **r =** $r = \frac{m^2}{\tilde{\sigma}^2 - m}$.

Therefore, $\tilde{r} = \frac{0.68^2}{0.81^2 - 0.68}$ = 3.56).

(iii) Pooling of centre-wise estimates: It is expected that the parameters of negative binomial distributions may not be same across the regions. For examples diffused nature of counterfeits may vary from the States to States. In such a case we have to resort to best available pooling method which is as under. Let for the i-th region (State/Union Territories. I = 1 to 35), X_i is the number of counterfeit notes per unit of a standard inspection checks in currency note inspection machine and the data get fit to negative binomial distribution with mean = u_i and variance ($U_i + U_i^2/r$). Therefore, if $X_i \sim RNBD(u_i, u_i + u_i^2/r_i)$ for I = 1 to 35, then the best linear unbiased estimator (BLUE) among the weighted average estimate is $\tilde{X} = \sum_{i=1}^{35} \frac{r_i X_i}{r_i u_i + u_i^2}$. In case diffusion indices are the same. i.e., $\phi_i = u_i / r_i = \ldots\ldots\ldots = \frac{1}{\pi} - 1$, we have, $\tilde{X} = \frac{1/\phi}{1+1/\phi} \sum_{i-1}^{35} \frac{r_i X_i}{r_i u_i + u_i^2} = \frac{\phi}{1+\phi} \sum_{i=1}^{35} \frac{X_i}{r_i}$ and $\frac{X_i}{r_i} \sim RNBD(\phi_i, \phi_i + \phi_i^2)$. Here classical model X_i ~NBD (r_i, π_i) is re-parameterized in terms of mean u=r_i $Ø_i$ and shape parameter r_i (number of successes before (n - r_i) failures in *n* number of trials).

2. Alternative estimation procedures[25]: The U.S. Treasury uses mainly two approaches namely the "parts-found-in-processing (PFP)" method and "the life-of-counterfeits (LOC)" method. The simplest PFP approach extrapolates the number of counterfeits per million found by the monetary authorities during currency processing to the entire stock of currency. PFP extends the approach to reflect the discovery of counterfeits outside the authorities' processing activities. In contrast, the LOC method extrapolates the flow of discovered counterfeits to the stock using estimates of the life of counterfeits in circulation. Bank of Canada attempted to adopt a revised method known as Chant's "composite approach" developed by Chant (2004). These methods with their limitations, relevance and applicability in the Indian context are described below.

[25] Sourced from "Counterfeiting: A Canadian Perspective" by John Chant, John (2004), Bank of Canada Working Paper 2004-33.

2.1. PFP approach: The simplest PFP approach estimates the number of circulating counterfeits of any denomination, C_D, as C_D = RBIPPM X $NIC_{D.}$ Here RBIPPM is the number of counterfeit notes detected per million notes processed by the central bank in the country and NIC_N is the outstanding stock of notes of denomination *N*. The PFP approach is somewhat simplistic and based on heuristics as the same would ideally have some correspondence to the actual stock of counterfeits only when, (i) detected counterfeits were found *only* during the central bank's processing activities, and (ii) the notes processed by the bank were representative of outstanding currency with respect to the share of counterfeits. In this case, the bank's detection rate for each denomination could be extrapolated to the stock of notes of that denomination to give an estimate of circulating counterfeits. It goes without saying that in a given period of time, all counterfeits floating in the system do not pass through central bank's note processing system. Thus, PFP method could give rise to a high degree of underestimation as and when public role in probable handling of counterfeits become significant. It is very much true for large currency holding by the people at large. This shortcoming of the PFP method of treating all counterfeits as if, they were detected in the central bank's note processing system, was subsequently adapted somewhat by the US Treasury to take into account the detections reported in other segments namely common public, banks or fake notes seized by police. The adapted version of PFP ($PFP^{/}$) adds the proportion of counterfeits detected by the public to the proportion detected during processing by the monetary authority as C_D = RBIPPM x s x NIC_D; $s=TD_D/BD_D = (PD_D+BD_D)/BD_D$ Here TD_D represents total detections of counterfeit notes of certain denomination N; PD_D, counterfeits (D denominated) held by the public; and BD_D, detections of denomination D made by the central bank/banking system. TD, PD, and BD are all measured as number of detections per year. The PFP approach represents a lower-bound estimate because it does not include the counterfeits detected outside the central bank. The $PFP^{/}$ approach represent a useful upper-bound estimate because it is based on the implausible assumption about the entire turnover of currency happening in private transactions so as to reveal probable extent of all the fake notes. However, though suffering from certain obvious limitations, it helps provide certain reporting of useful numbers, which when analysed in a disaggregated manner over a period of time, might provide a clue to certain dimensions of counterfeit detection ability in the system. It is of course argued in recent analyses of US FED and Treasury Office that it is well nigh impossible for bulk of counterfeit US dollar currency to remain in circulation without getting intercepted by the banking system or law enforcing machineries.

2.2. Life-of-counterfeit approach: Stock of circulating counterfeits can also be estimated using the "life-of counterfeit" (LOC) method. This method extrapolates the flow of discovered counterfeits to the total stock by using the estimated life of counterfeits. With this approach, the number of circulating counterfeits of certain denomination (D) is $C_D = LOC_D \times TD_D$ where, LOC_D represents the life of counterfeits, and TD_D is the annual recovery of counterfeits of denomination D. The shortcoming of the LOC approach is that past history on the circulating life of counterfeits are meagre. Of course, by putting the year of printing in the genuine currency notes, data base on life of different denominations of currency notes issued in different series now would enable one to estimate average life of currency notes, which could proxy for similarly counterfeited notes.

2.3. The composite method: The composite method (COMP) combines elements of both PFP and LOC to estimate the stock of circulating counterfeits. It draws on the LOC approach by using the information on the life of counterfeit notes. It then uses PFP, together with data on the public's detection of counterfeits, to anchor estimates of the counterfeit stock on assumptions about the public efficiency in detecting counterfeits. The COMP method uses more data for its estimates than either the LOC or PFP approaches. These data include information about the life of counterfeits, the rate at which counterfeits are detected by the monetary authority during processing, and the annual flow of counterfeits detected outside the banking system. This approach explicitly recognizes that screening for counterfeits takes place both inside and outside the central bank. The public and financial institutions, in their transactions and processing of currency, are the sources of screening outside the monetary authority. The efficiency of screening when currency is transferred among individuals, businesses, and financial institutions indicates the proportion of counterfeits that originally existed in the batches of currency before they were sent to the central bank. The COMP method estimates the stock of outstanding counterfeits using three separate elements. **(a)** Any batch of currency processed by the central bank first turns over in a private sector transaction, where '*e'* of the counterfeits are detected before it is passed to the central bank, or where the remaining counterfeit notes are detected. If PPM (parts Per Million) is the original proportion of counterfeits in circulating currency, then PPM is described as PPM = RBIPPM/ (1-e) (0<e<1). Here RBIPPM is the proportion of counterfeits detected by the Reserve Bank. The first element expresses the relation between the stock of outstanding counterfeits, *C*, of any denomination (D) and detections of counterfeits of that denomination, given the assumed efficiency of public screening, e, and the proportion of counterfeiting detected by central bank (RBIPPM), $C(e)_D = PPM \times NIC_D = RBIPPM \times NIC_D/(1-e)$.

It builds on the PFP method by allowing for different efficiencies of public detection. **(b)** The second element deals with the turnover of currency needed to account for the actual level of public detection of counterfeits during a year, given the efficiency of public screening. The estimated turnover, T of counterfeits of any denomination D is given by $T\ (e)_D = PD_D/(e \times PPM \times NIC_D)$. Here, PD_D is the detection of denomination D made by public per year. Here the denominator measures public detections per turnover of the circulating stock of denomination D. **(c)** The third element is estimated life of counterfeit notes, which is given by $LOC_D = C\ (e)_D/TD_D$.

Data are readily available for *RBIPPM*, the proportion of counterfeiting detected by central bank and NIC_D. Each equation, however, requires information on unknowns in order to estimate $C(e)_D$. The unknowns are *e*, LOC_D and $T\ (e)_D$. Values for T_D and LOC_D could be derived using knowledge about the turnover rate of the currency or the life of counterfeits.

2.4 Method for estimating life of counterfeits: Average life of counterfeits can be estimated using the recovery data of high quality counterfeit notes circulating at different time points. The rate of decay of the stock of counterfeits may be derived as follows. The stock of counterfeits at any time *t* periods after the series ceased to be introduced, C_t, can be represented as $C_t = C_0.Exp(-dt)$, where C_0 is the where Co is the stock at the time new counterfeits ceased to be introduced, and d is the rate of decay of the counterfeit stock. But since the rate of decay, $r_t = d.C_t$; $r_t = r_0.Exp(-dt)$. Thus, the decay rate of circulating counterfeits can be estimated by the equation: $\ln r = \ln r_0 - dt$. The lifespan of counterfeits of other denominations may be obtained using the lifespan data of notes of that denomination. These data, together with the assumption that turnover and currency life are inversely proportional, give estimates of the turnover rates for each denomination. The estimated turnover rates are substituted into the second equation in the above set of five equations to generate relevant estimates of '*e*' for each denomination. To avoid complexity, one can start with some hypothesized estimate of LOC, which may periodically be checked with empirical evidences and judgment, based on newly configured data base on year of issuances and recording the same from counterfeits detected in the recent period.