International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
SJIF (2024): 6.623

Predicting a Random Determinant with i.i.d. Bernoulli Variates

Shashi Kant Agrawal[1], Pinky Pandey[2], Soubhik Chakraborty[3]

[1]University Department of Mathematics, Ranchi University, Ranchi – 834008, India
Corresponding Author Email: *shashikantagrawal658[at]gmail.com*

[2]Department of Mathematics, Nirmala College, Ranchi University, Ranchi – 834002, India
Email: *unix.pinky[at]gmail.com*

[3]Department of Mathematics, Birla Institute of Technology Mesra, Ranchi – 835215, India
Email: *soubhikc[at]yahoo.co.in*

Abstract: ***This paper gives a probabilistic analysis of a determinant D of order 2 and 3 in which the elements are i.i.d. Bernoulli variates. Using Chebyshev's inequality, fiducial limits of D are obtained for order 2 and 3. The results may be compared with those obtained for other standard probability distributions.***

Keywords: Random Determinant; Bernoulli Distribution; Chebyshev's inequality

Mathematics Subject Classification: 62P99

1. Introduction

Bernoulli distributions find applications in reliability and survival analysis. Bernoulli distribution is used in medicine and clinical trials to model the success rate of a certain drug or the outcomes of a clinical trial. For example, when developing a new drug, pharmaceutical scientists can use Bernoulli distributions to calculate the probability that a person will be cured or not cured with the help of new drug. Bernoulli distributions are not only useful for mathematicians and statisticians; they also have a crucial role to play in data analytics, data science, and machine learning. These applications motivated us to study a random determinant filled with i.i.d. (independently and identically distributed) Bernoulli variates. We shall be using Chebyshev's inequality to predict such a random determinant. Chebyshev's inequality states that for a random variable X, $P\{E(X) - kSD(X) < X < E(X) + kSD(X)\} \geq 1-\frac{1}{k^2}$, for some constant k which is usually some positive integer. Here E(X) is the mathematical expectation of X and SD(X) gives the standard deviation of X. The proof of Chebyshev's inequality can be found in any standard text on statistics (e.g Gupta and Kapoor, 2014). See also Solary (2018). For a sound literature on determinants, see Muir (2011).

2. Previous Work

Surprisingly there is not much work in this problem except for the paper by Wise and Hall (1991) in which the distribution of a determinant of order 2 for i.i.d U(0,1) variates has been given. Saha and Chakraborty (2019) have extended the study for i.i.d U(0, θ) variates and i.i.d. U(1, 2, 3,….t) variates and, using Chebyshev's inequality, obtained the fiducial limits of D for order 2 and 3 for several other probability distributions. See also Saha and Chakraborty (2020).

3. Bernoulli Distribution

A random variable X is said to have a Bernoulli distribution with parameter p if its p.m.f is given by:

$$P(X = x) = \begin{cases} p^x(1-p)^{1-x} & ; \quad \text{for } x = 0,1 \\ 0 & ; \quad \text{otherwise} \end{cases}$$

The parameter p satisfies $0 \leq p \leq 1$. Often $(1-p)$ is denoted as q.

A random experiment whose outcomes are of two types, success S and failure F, occurring with probabilities p and q respectively, is called a Bernoulli trial. If for this experiment, a random variable X is defined such that it takes value 1 when S occurs and 0 if F occurs, then X follows a Bernoulli distribution.

4. Our Contribution

First, consider a determinant D of order 2 as $D = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = (a_{11}a_{22} - a_{12}a_{21})$

Let a_{ij}'s be i.i.d Bernoulli variates.

Remark: Mean of Bernoulli Distribution E(X) = p ……(i)
and Variance of Bernoulli Distribution V(X) = pq ….(ii)

For proof, see Gupta and Kapoor (2014).

Then, E(D)
$= E(a_{11}a_{22} - a_{12}a_{21})$
$= E(a_{11}a_{22}) - E(a_{12}a_{21})$
$= E(a_{11})E(a_{22}) - E(a_{12})E(a_{21})$ (Since, a_{ij}'s are independent)
$= p \cdot p - p \cdot p$
$= p^2 - p^2$
$= 0$

Now,

We know that
$V(X) = E(X^2) - \{E(X)\}^2$
$\Rightarrow E(X^2) = V(X) + \{E(X)\}^2$
$= pq + p^2$ …………………using (i) and (ii)
$= p(q + p)$
$= p$ (since $p + q = 1$)
$\therefore E(X^2) = p$ …………(iii)
next

$\therefore \text{Var}(D) = \text{Var}(a_{11}a_{22} - a_{12}a_{21})$
$= E(a_{11}a_{22} - a_{12}a_{21})^2 - \{E(a_{11}a_{22} - a_{12}a_{21})\}^2$
$\Rightarrow \text{Var}(D) = E(a_{11}a_{22} - a_{12}a_{21})^2 - 0$
$\Rightarrow \text{Var}(D) = E(a_{11}a_{22} - a_{12}a_{21})^2$
$\Rightarrow \text{Var}(D) = E(a_{11}^2)\, E(a_{22}^2) + E(a_{12}^2)\, E(a_{21}^2) - 2E(a_{11})\, E(a_{22})\, E(a_{12})\, E(a_{21})$
($\because a_{ij}$'s are independent)

$\therefore \text{Var}(D) = E(X^2)\, E(X^2) + E(X^2)\, E(X^2) - 2\, E(X)\, E(X)\, E(X)\, E(X)$
$= 2\{E(X^2)\}^2 - 2\{E(X)\}^4$
$= 2p^2 - 2p^4$ ………..using (i) and (iii)
$= 2p^2(1 - p^2)$
$= 2p^2(1 + p)(1 - p)$
$= 2p^2q(1 + p)$

$\therefore$ Standard Deviation (D)
$\sigma = \{2p^2q(1 + p)\}^{\frac{1}{2}}$
$\Rightarrow \sigma = p\sqrt{2q(1 + p)}$

$\therefore$ using Chebyshev's inequality yields
$P(E(D) - k\, SD(D) < D < E(D) + k\, SD(D)) \geq \left(1 - \frac{1}{k^2}\right)$
$\Rightarrow P(-k\sigma < D < k\sigma) \geq \left(1 - \frac{1}{k^2}\right)$
Where, $\sigma = p\sqrt{2q(1 + p)}$

Next, assume that D is of order 3.

Consider $D = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$
$= a(ei - fh) - b(di - fg) + c(dh - ge)$
$= aei - afh - bdi + bfg + cdh - cge$
where a, b, c………..i are i.i.d. Bernoulli variates.

Proceeding similarly as in the above case when D was of order 2, we get
$E(D) = E(aei - afh - bdi + bfg + cdh - cge)$
$= E(aei) - E(afh) - E(bdi) + E(bfg) + E(cdh) - E(cge)$

Let $E(a_{ij}) = p$

Then
$E(D) = p^3 - p^3 - p^3 + p^3 + p^3 - p^3$
$\Rightarrow E(D) = 0$ ……………………..(iv)

So,
$V(X) = E(X^2) - \{E(X)\}^2$
$\therefore V(D) = E(D^2) - \{E(D)\}^2$
$= E\{(aei - afh - bdi + bfg + cdh - cge)^2\} - 0$ ……………… using (iv)
$= E(aei - afh - bdi + bfg + cdh - cge)^2$
$= E[(aei - afh - bdi + bfg + cdh - cge)(aei - afh - bdi + bfg + cdh - cge)]$
$= E[a^2e^2i^2 - a^2efhi - abdei^2 + abefgi + acdehi - ace^2gi - a^2efhi + a^2f^2h^2 + abdfhi - abf^2gh - acdfh^2 + acefgh - abdei^2 + abdfhi + b^2d^2i^2 - b^2dfgi - bcd^2hi + bcdegi + abefgi - abf^2gh - b^2dfgi + b^2f^2g^2 + bcdfgh - bcefg^2 + acdehi - acdfh^2 - bcd^2hi + bcdfgh + c^2d^2h^2 - c^2degh - ace^2gi + acefgh + bcdegi - bcefg^2 - c^2degh + c^2e^2g^2]$
$= (E(X^2))^3 - E(X^2)(E(X))^4 - E(X^2)(E(X))^4 + (E(X))^6 + (E(X))^6 - E(X^2)(E(X))^4 - E(X^2)(E(X))^4 + (E(X^2))^3 + (E(X))^6 - E(X^2)(E(X))^4 - E(X^2)(E(X))^4 + (E(X))^6 - E(X^2)(E(X))^4 + (E(X))^6 + (E(X^2))^3 - E(X^2)(E(X))^4 - E(X^2)(E(X))^4 + (E(X))^6 + (E(X))^6 - E(X^2)(E(X))^4 - E(X^2)(E(X))^4 + (E(X^2))^3 + (E(X))^6 - E(X^2)(E(X))^4 + (E(X))^6 - E(X^2)(E(X))^4 - E(X^2)(E(X))^4 + (E(X))^6 + (E(X^2))^3 - E(X^2)(E(X))^4 - E(X^2)(E(X))^4 + (E(X))^6 + (E(X))^6 - E(X^2)(E(X))^4 - E(X^2)(E(X))^4 + (E(X^2))^3$

Let $(E(X))^6 = A$
$E(X^2)(E(X))^4 = B$
$(E(X^2))^3 = C$

Then
$\text{Var}(D) = C - B - B + A + A - B - B + C + A - B - B + A - B + A + C - B - B + A + A - B - B + C + A - B + A - B - B + A + C - B - B + A + A - B - B + C$

$\text{Var}(D) = 12A - 18B + 6C$ ……………(v)

$\because A = (E(X))^6$
$= p^6$
$B = E(X^2)(E(X))^4$
$= p \cdot p^4$
$= p^5$
and
$C = (E(X^2))^3$
$= p^3$

From equation (v)
$\text{Var}(D) = 12A - 18B + 6C$
$\therefore \text{Var}(D) = 12p^6 - 18p^5 + 6p^3$
$= 6p^3(2p^3 - 3p^2 + 1)$
$= 6p^3(p - 1)^2(2p + 1)$
$= 6p^3(-q)^2(2p + 1)$
$\text{Var}(D) = 6p^3q^2(1 + 2p)$

$\therefore$ Standard Deviation(D)
$\sigma = \{6p^3q^2(1 + 2p)\}^{\frac{1}{2}}$
$\Rightarrow \sigma = pq\sqrt{6p(1 + 2p)}$

$\therefore$ using Chebyshev's inequality yields
$P(E(D) - k\, SD(D) < D < E(D) + k\, SD(D)) \geq \left(1 - \frac{1}{k^2}\right)$
$\Rightarrow P(-k\sigma < D < k\sigma) \geq \left(1 - \frac{1}{k^2}\right)$

where $\sigma = pq\sqrt{6p(1+2p)}$

5. Discussion

It is of interest to compare the results on fiducial limits of D obtained for Bernoulli Distribution as inputs with those obtained for other probability distributions. We refer the reader to tables 1 and 2 of Saha and Chakraborty (2020).

6. Conclusion

1) If D is of order 2 with i.i.d Bernoulli variates as input,
   $P(-k\sigma < D < k\sigma) \geq \left(1 - \frac{1}{k^2}\right)$
   where $\sigma = p\sqrt{2q(1+p)}$
   with E(D) = 0 and Var(D) = $2p^2q(1+p)$

2) If D is of order 3 with i.i.d Bernoulli variates as input,
   $P(-k\sigma < D < k\sigma) \geq \left(1 - \frac{1}{k^2}\right)$
   Where, $\sigma = pq\sqrt{6p(1+2p)}$
   with E(D) = 0 and Var(D) = $6p^3q^2(1+2p)$

Ethical Statement: The authors hereby declare that this research did not receive any funding. They further declare that they do not have any conflict of interest.

References

[1] Wise, G. L. and Hall, E. B., A note on the distribution of the determinant of a random matrix. *Statistics and Probability Letters* 11 (1991) 147- 148
[2] Muir, T., A Treatise on the theory of determinants, Nabu Press, 2011
[3] Gupta, S.C. and Kapoor, V. K., *Fundamental of Mathematical Statistics,* Sultan Chand and Sons, 2014
[4] Solary, M.S. (2018) Converting a Chebyshev polynomial to an ordinary polynomial with series of powers form, *Journal of Interdisciplinary Mathematics,* 21:7-8, 1519-1532, DOI:10.1080/0972502.2017.1303949
[5] Saha, N and Chakraborty, S, using Chebyshev's inequality to predict a Random Determinant for i.i.d Gamma and Weibull distributions, Journal of Statistics and Mangement Systems, vol.24, No.3, 2020, P. 613 - 623
[6] Saha. N. and Chakraborty, S., Probabilistic Analysis of a Random Determinant, GRIN Verlag, 2019