年賀状の番号の分布

2018.1.1
渡邉 俊夫

年賀状のお年玉くじ

年賀状にはお年玉くじが付いていて、6桁の数字が当選番号に一致すると賞品をもらうことができる。１等の当選番号は6桁で1つ（100万本に1本）、２等は下4桁で1つ（1万本に1本）、３等の当選番号は下2桁で2つ（100本に2本）である。

毎年100枚くらいの年賀状が届く場合、１等や2等が当たる可能性は低く、当たる可能性が高いのは3等（賞品は年賀切手シート）である。

３等の本数からすれば、届いた100枚の年賀状の中に、2つの当選番号がそれぞれ1枚ずつありそうだが、それは正しいだろうか？

年賀状の当選番号と枚数(実測)

過去8年(ただし、2012年は実父、2014年は祖父の喪中のため除いた)の当選番号と当選枚数を下表に示す。

年		干支	当選番号		届いた枚数	当選枚数		当選計
2008年	平成20年	子	37	64	95	2	0	2
2009年	平成21年	丑	46	94	104	1	1	2
2010年	平成22年	寅	00	52	95	0	2	2
2011年	平成23年	卯	02	69	104	0	2	2
2013年	平成25年	巳	29	70	102	0	1	1
2015年	平成27年	未	27	30	98	0	1	1
2016年	平成28年	申	69	90	96	0	1	1
2017年	平成29年	酉	45	51	95	0	1	1

年賀状の当選番号と枚数（実測）

前頁の表に示したように、年2枚当たったのが8年中4回、年1枚当たったのが8年中4回だった。また、2枚当たった年も、2つの当選番号が1枚ずつ当たったのは1回だけで、それ以外の年は一方の当選番号が2枚、他の当選番号は0枚だった。

これは、年賀状の番号がランダムに届くとすると、定性的には妥当である。なぜなら、ランダムならば、ある1枚の番号は他の番号とは無関係に決まるからである。100枚の年賀状が届くとき、00から99までの100通りの番号が均等に1枚ずつ届くのは、むしろランダムではない。したがって、一方の当選番号が0枚でも不思議ではない。

むろん、実例からもわかるように、当選番号が0枚の場合もあるが、2枚以上の場合もある。年賀状が100枚届くときの当選枚数の期待値は（当選番号が2つあるので）2枚になる。

二項分布

0 から $n-1$ までの番号（3等の場合は $n=100$）のうちの1つが書かれた年賀状をランダムに m 枚受け取ったとき、同じ番号の年賀状が何枚あるかを考える。

ある番号のカードを k 枚（$0 \le k \le m$）受け取る確率は、m 枚のうち k 枚がその番号で、残りの $m-k$ 枚がそれ以外の番号になることから

$$p(k) = {}_{m}C_{k}\left(\frac{1}{n}\right)^{k}\left(1-\frac{1}{n}\right)^{m-k} = \frac{m!}{(m-k)!\,k!}\frac{1}{n^k}\frac{(n-1)^{m-k}}{n^{m-k}}$$

$$= \frac{m!}{(m-k)!\,k!}\frac{(n-1)^{m-k}}{n^{m}}$$

であり、これは二項分布である。

ポアソン分布

二項分布

$$p(k) = \frac{m!}{(m-k)!\,k!}\left(\frac{1}{n}\right)^k\left(1-\frac{1}{n}\right)^{m-k} = \frac{m!}{(m-k)!\,k!}\frac{(n-1)^{m-k}}{n^m}$$

は、$n \gg 1$ のとき、$m \gg k$ となるから、$\mu = m/n$ とおくと

$$p(k) = \frac{m(m-1)\cdots(m-k+1)}{n^k}\frac{1}{k!}\left(1-\frac{1}{n}\right)^{m-k}$$

$$= \frac{m}{n}\frac{m-1}{n}\cdots\frac{m-k+1}{n}\frac{1}{k!}\left(1-\frac{1}{n}\right)^{m-k} \cong \left(\frac{m}{n}\right)^k\frac{1}{k!}\left(1-\frac{1}{n}\right)^m$$

$$= \frac{\mu^k}{k!}\left(1-\frac{\mu}{m}\right)^m \cong \frac{\mu^k}{k!}e^{-\mu}$$

と近似できる。これは平均値 μ のポアソン分布である。

二項分布とポアソン分布

二項分布

$$p(k) = \frac{m!}{(m-k)!\,k!}\frac{(n-1)^{m-k}}{n^m}$$

および、ポアソン分布($\mu = m/n$)

$$p(k) = \frac{\mu^k}{k!}e^{-\mu}$$

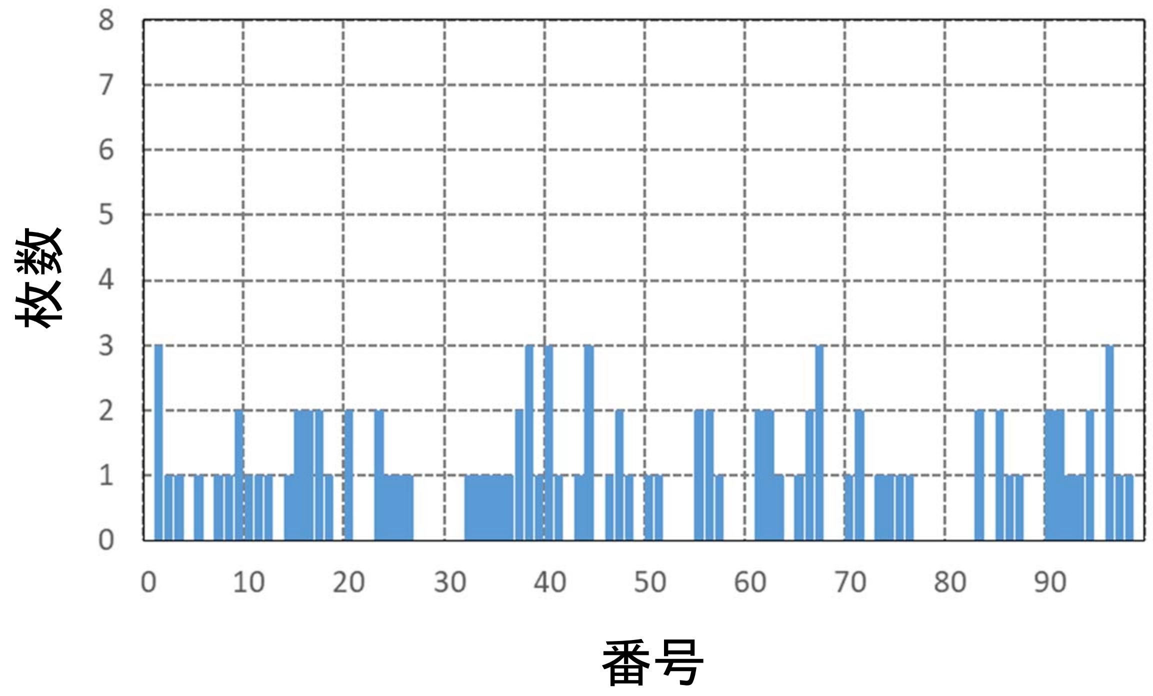
において、$n = 100, m = 100$ のときの $p(k)$ の値を右表に示す。両者は、ほぼ一致している。

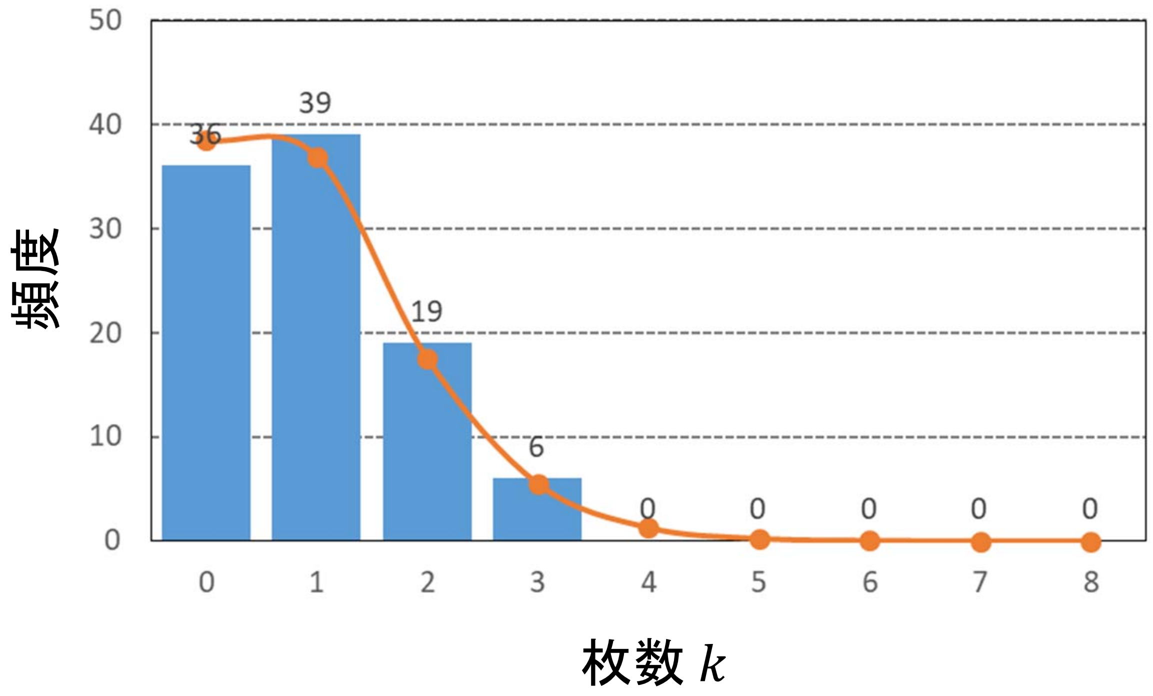
$p(0) \approx 0.37$ であるから、100通りの番号のうち約37通りは0枚である。

k	二項分布 ($n = m = 100$)	ポアソン分布 ($\mu = 1$)
0	0.3660	0.3679
1	0.3697	0.3679
2	0.1849	0.1839
3	0.0610	0.0613
4	0.0149	0.0153
5	0.0029	0.0031
6	4.6×10^{-4}	5.1×10^{-4}
7	6.3×10^{-5}	7.3×10^{-5}
8	7.4×10^{-6}	9.1×10^{-6}

年賀状の番号と枚数（実測）：2008年

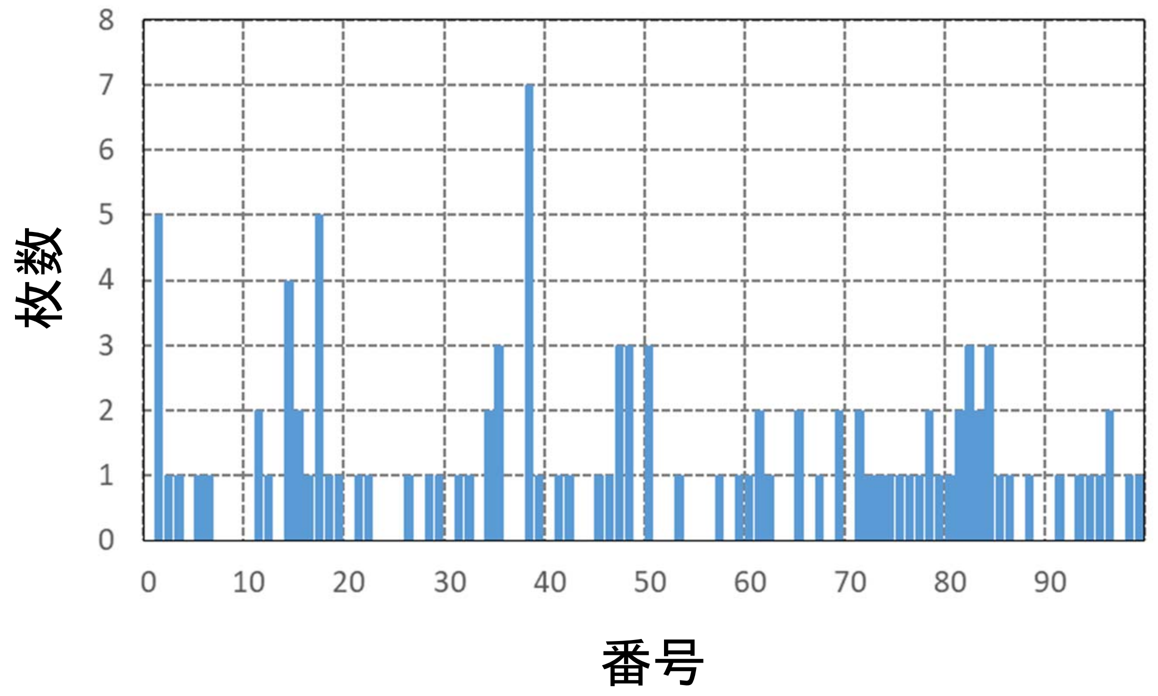
2008年に届いた年賀状（$m = 95$）の下2桁の各番号の枚数を下左図に示す。また、下右図の棒グラフはその番号の枚数が k 枚の頻度を示す。赤実線は二項分布による計算結果である。

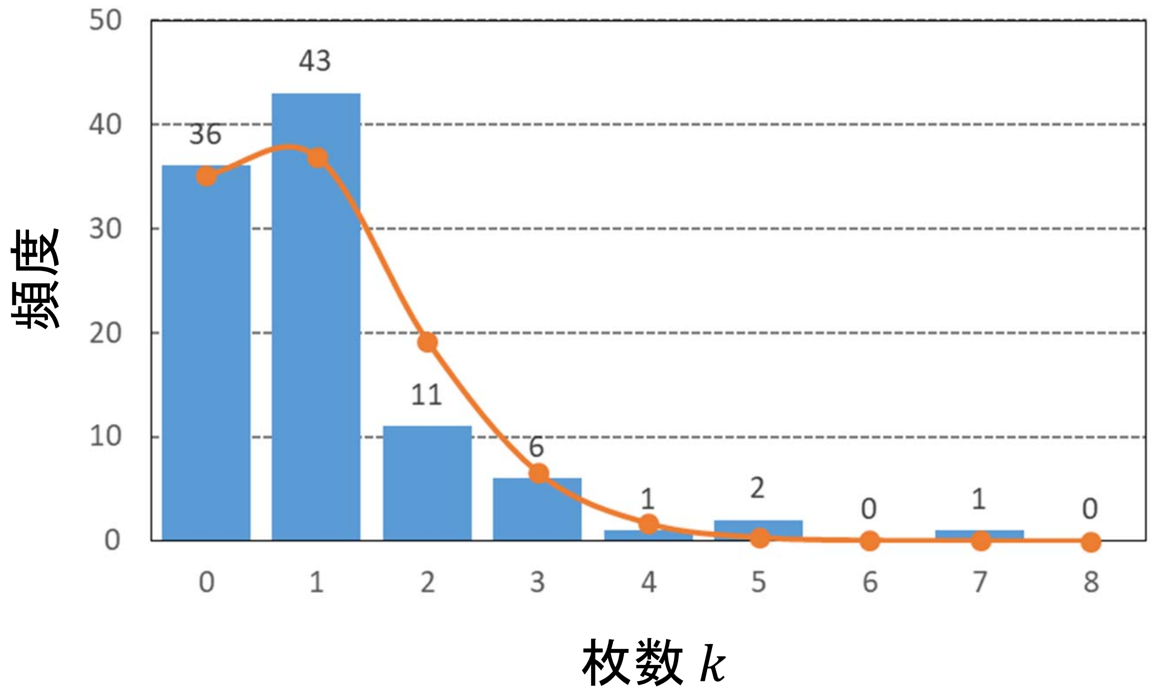




年賀状の番号と枚数（実測）：2009年

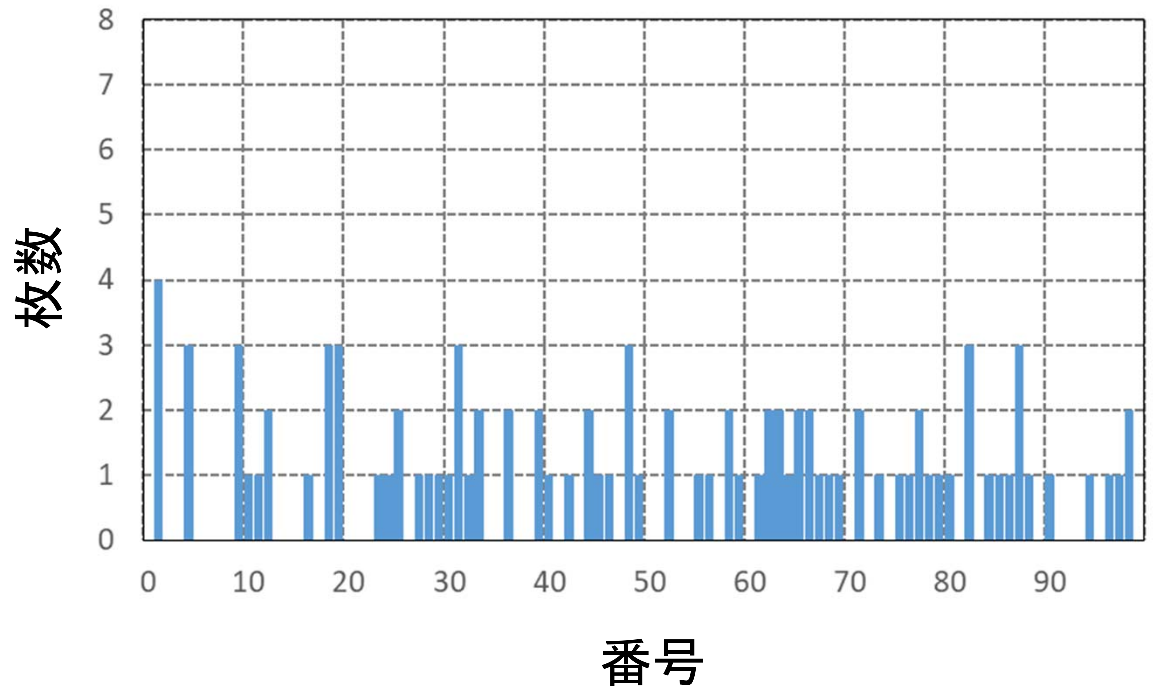
2009年に届いた年賀状（$m = 104$）の下2桁の各番号の枚数を下左図に示す。また、下右図の棒グラフはその番号の枚数が k 枚の頻度を示す。赤実線は二項分布による計算結果である。

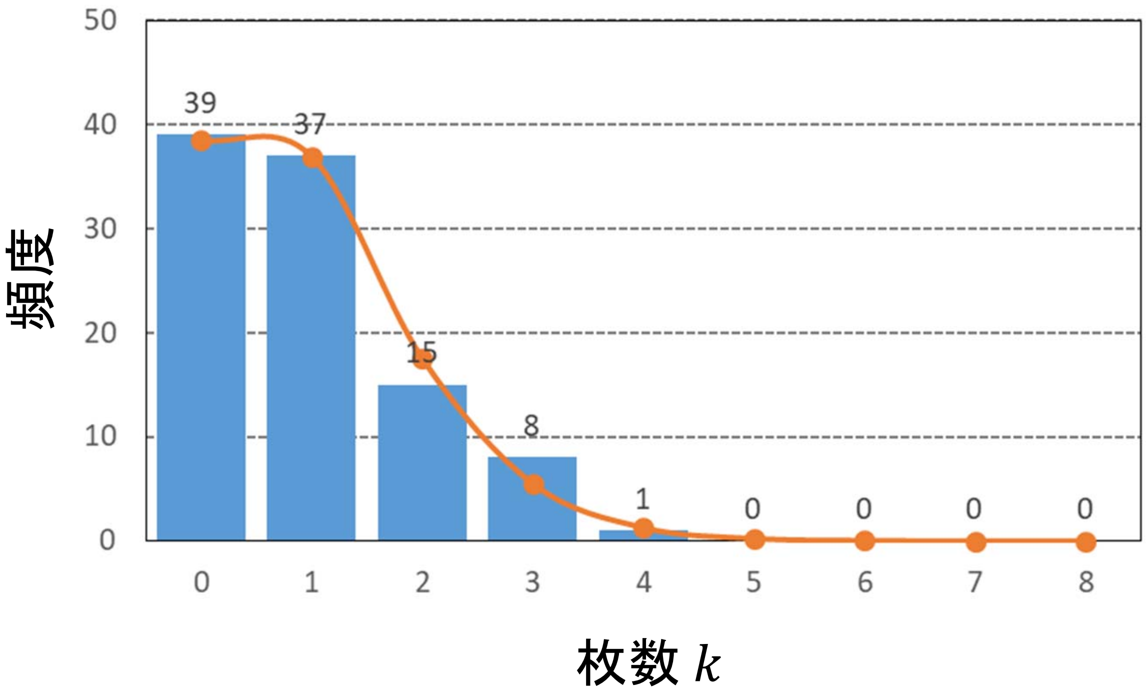




年賀状の番号と枚数(実測):2010年

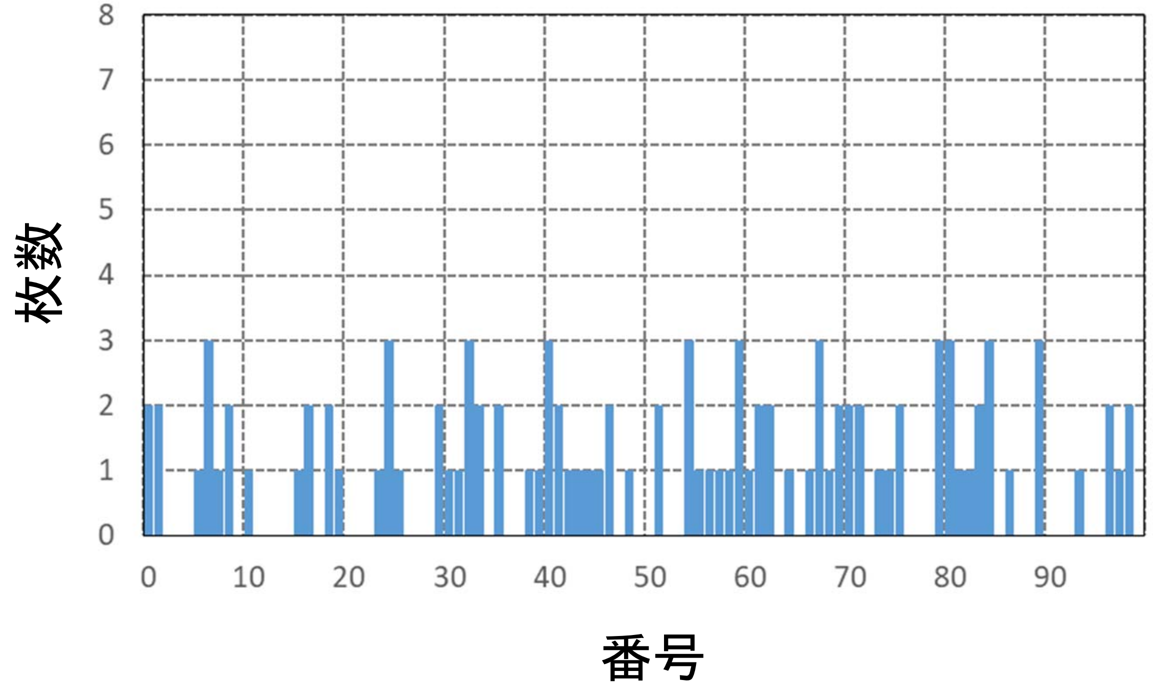
2010年に届いた年賀状($m = 95$)の下2桁の各番号の枚数を下左図に示す。また、下右図の棒グラフはその番号の枚数が k 枚の頻度を示す。赤実線は二項分布による計算結果である。

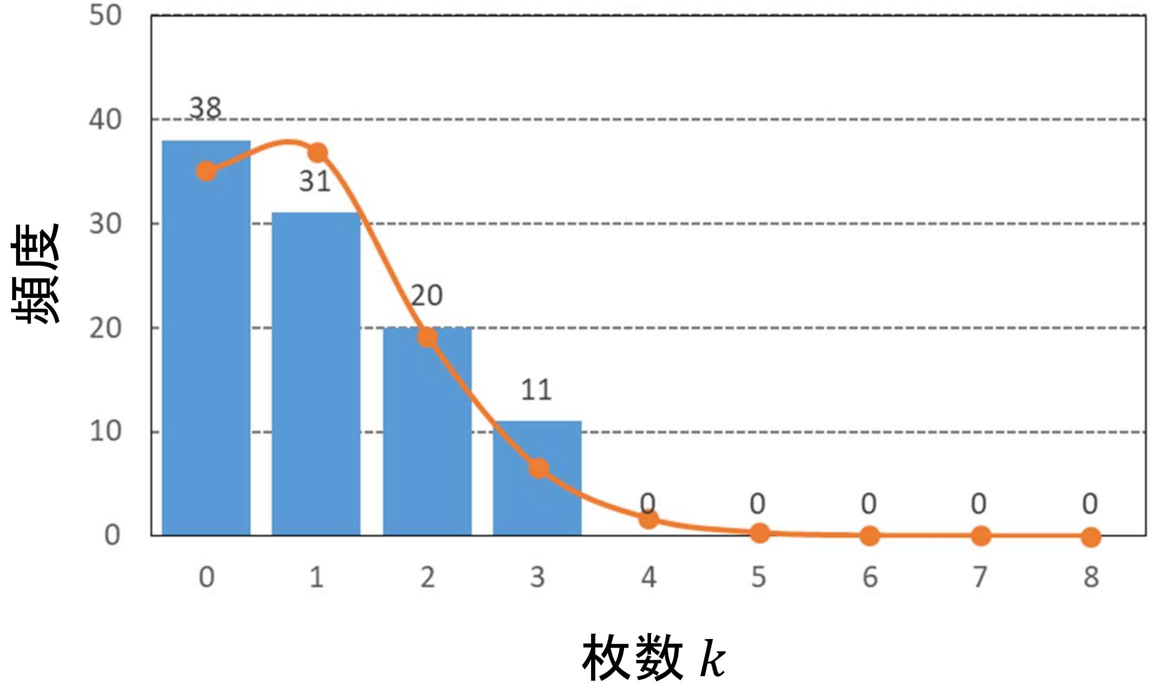




年賀状の番号と枚数（実測）：2011年

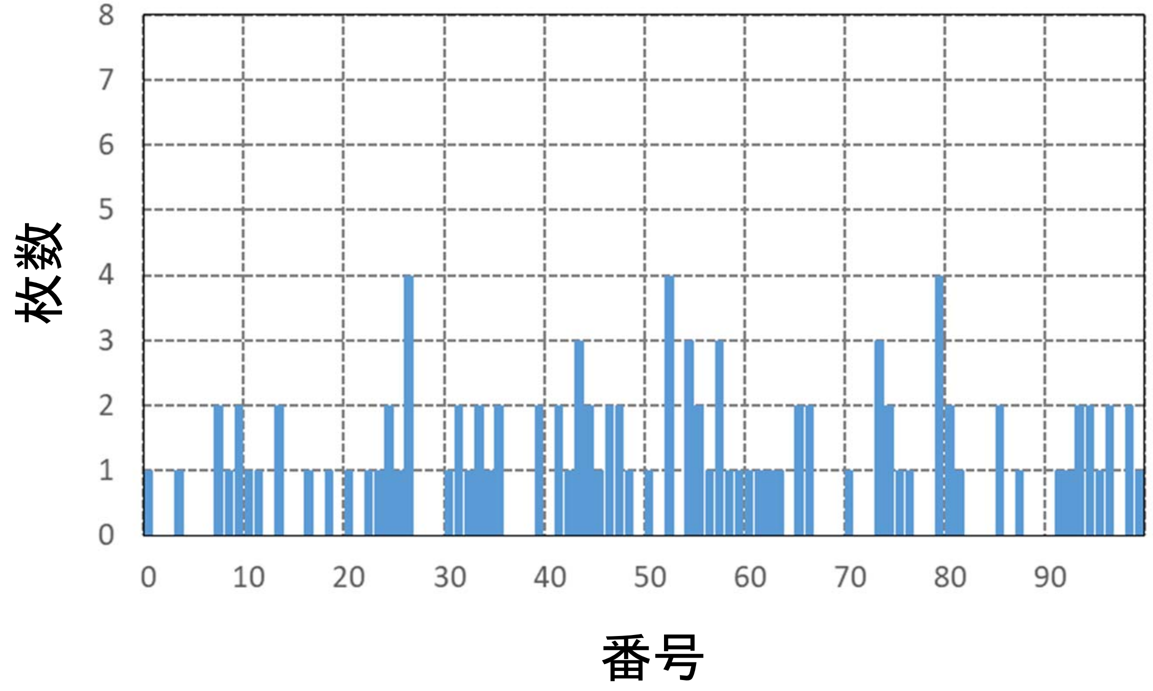
2011年に届いた年賀状（$m = 104$）の下2桁の各番号の枚数を下左図に示す。また、下右図の棒グラフはその番号の枚数が k 枚の頻度を示す。赤実線は二項分布による計算結果である。

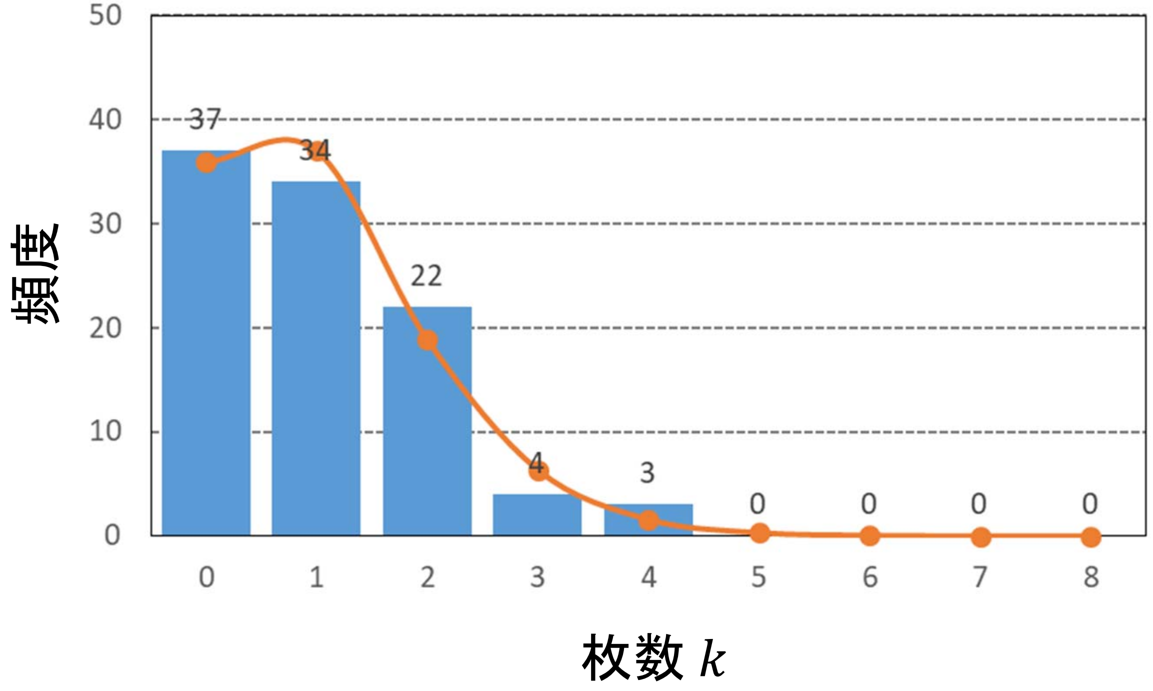




年賀状の番号と枚数(実測):2013年

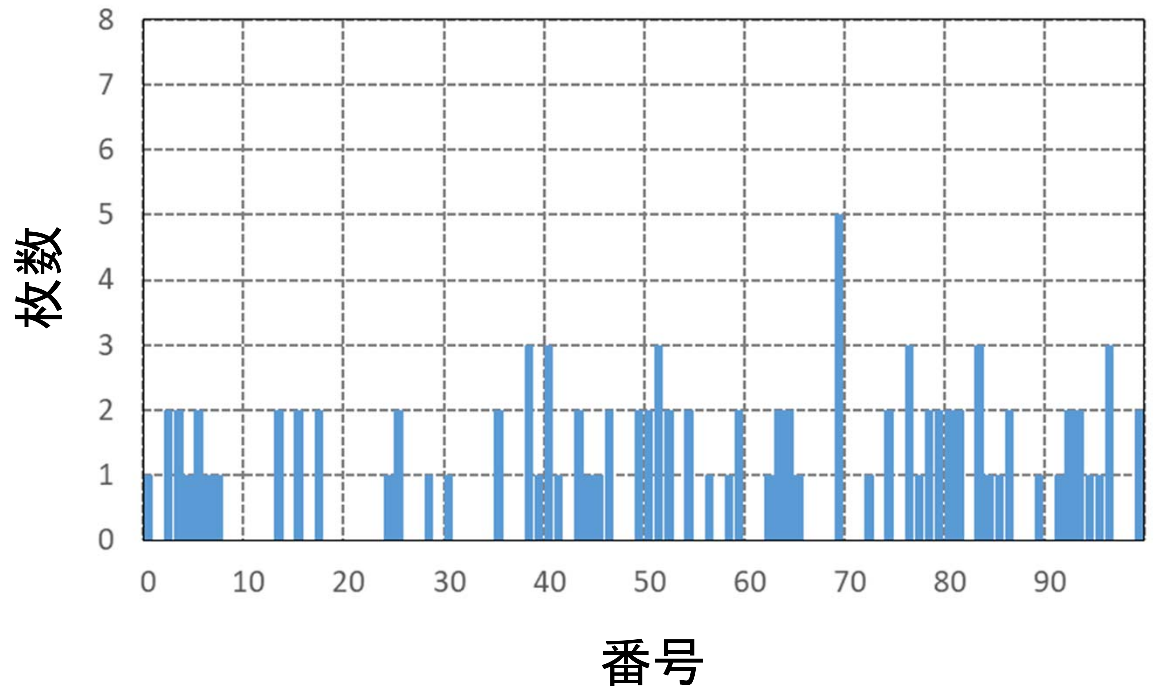
2013年に届いた年賀状($m = 102$)の下2桁の各番号の枚数を下左図に示す。また、下右図の棒グラフはその番号の枚数が k 枚の頻度を示す。赤実線は二項分布による計算結果である。

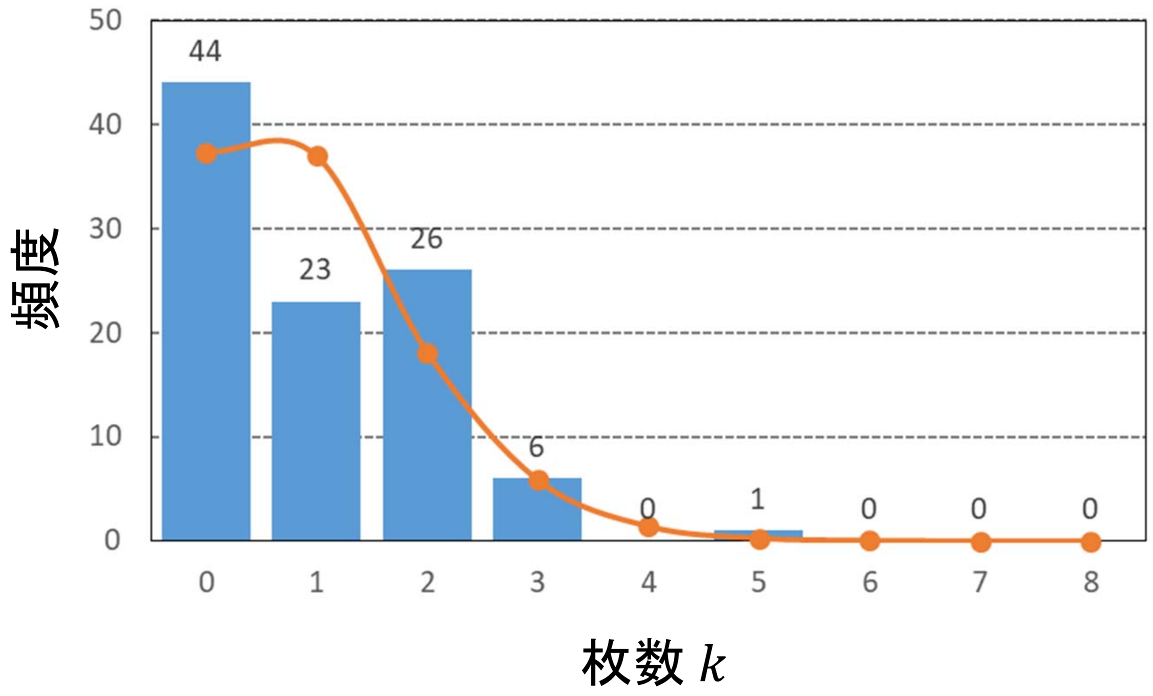




年賀状の番号と枚数(実測):2015年

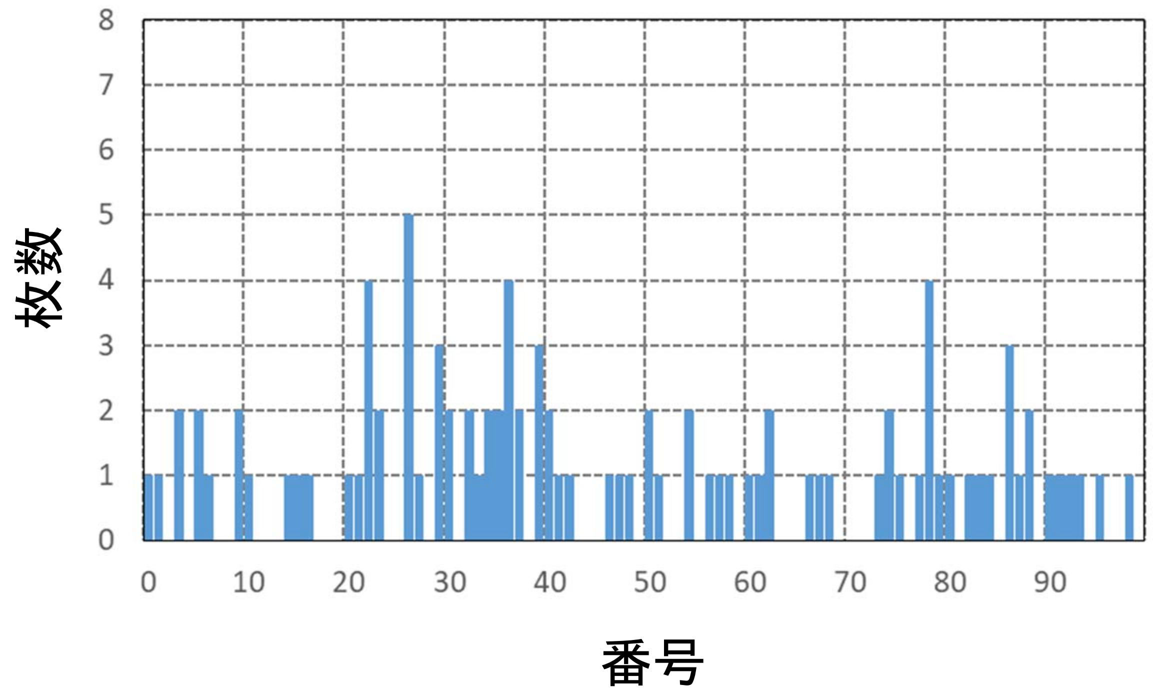
2015年に届いた年賀状($m = 98$)の下2桁の各番号の枚数を下左図に示す。また、下右図の棒グラフはその番号の枚数が k 枚の頻度を示す。赤実線は二項分布による計算結果である。

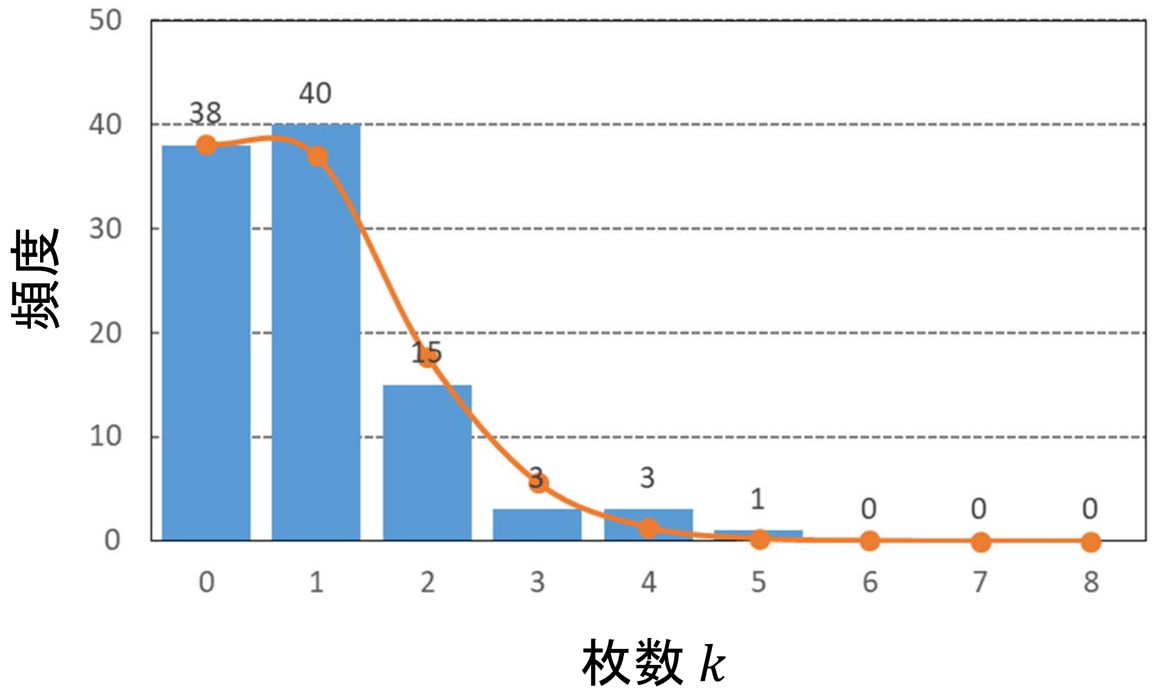




年賀状の番号と枚数（実測）：2016年

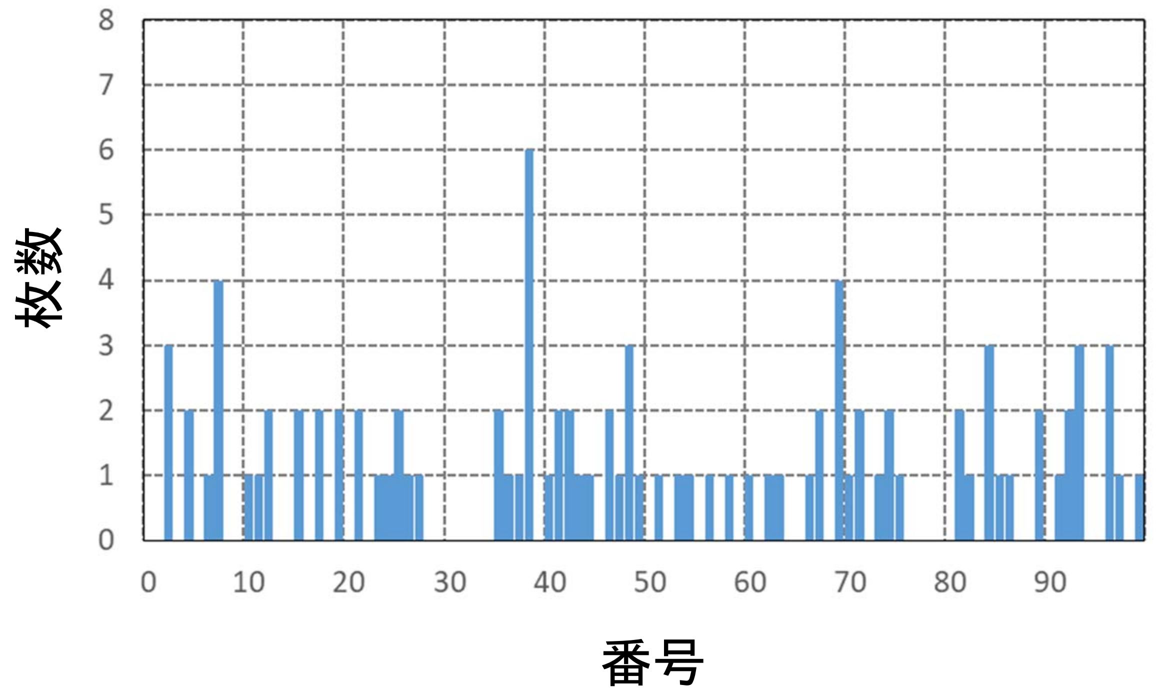
2016年に届いた年賀状（$m = 96$）の下2桁の各番号の枚数を下左図に示す。また、下右図の棒グラフはその番号の枚数が k 枚の頻度を示す。赤実線は二項分布による計算結果である。

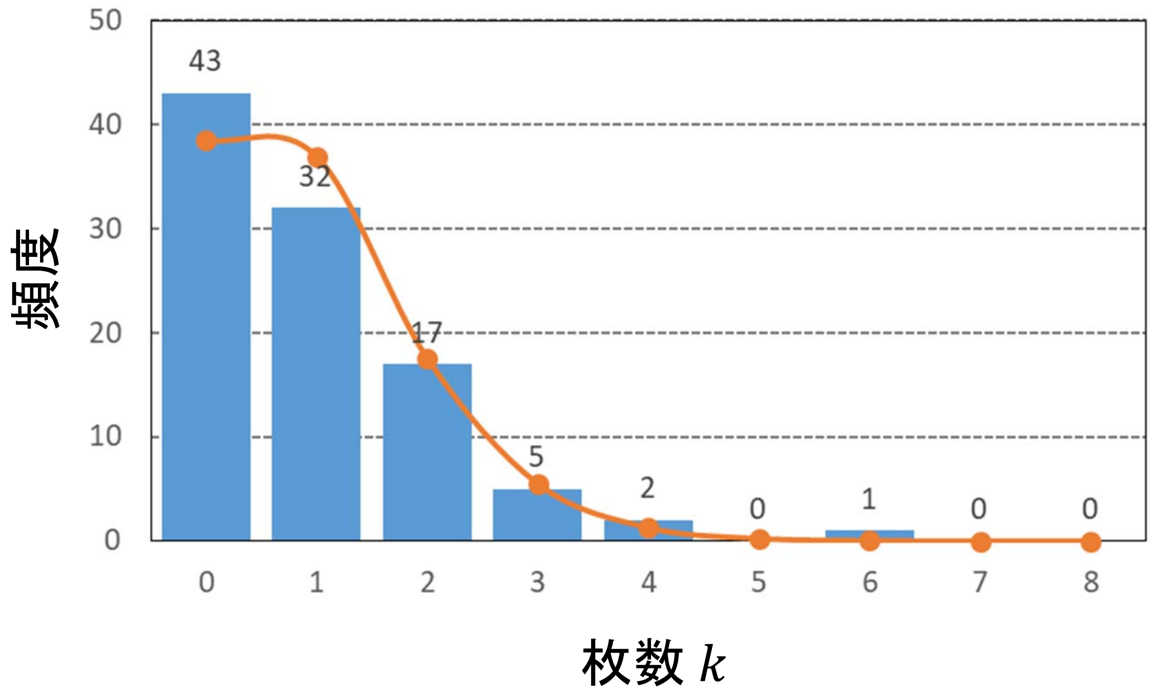




年賀状の番号と枚数(実測):2017年

2017年に届いた年賀状($m = 95$)の下2桁の各番号の枚数を下左図に示す。また、下右図の棒グラフはその番号の枚数が k 枚の頻度を示す。赤実線は二項分布による計算結果である。





まとめ

自分宛てに届く年賀状のお年玉くじの下2桁の番号は、00～99の100通りが均等には分布していない。

年賀状の番号がランダムに届くとすれば、その分布は二項分布（ポアソン分布）になる。

100枚の年賀状が届く場合、下2桁の番号100通りのうち約37通りは、その番号の枚数が0枚と見積もられる。1つの当選番号だけについて言えば、当選枚数の期待値は1枚であるが、当たる確率は約63%である。

謝辞

いつも年賀状をいただく皆様に心より感謝申し上げます。