Probability Analysis of the "49ers" Sample

Prepared by

Dr Paul Yip, PhD

Senior Lecturer,

Department of Statistics and Actuarial Science

The University of Hong Kong.

Hong Kong

Background:

I have been given to know that -

- In July 2001 Cathay Pacific Airways (CPA) terminated the contracts of service of 50 pilots.
- The group became known as "the 49ers" as 49 of the pilots were terminated on one day (9^{th} July 2001)
- 49 of the 50 pilots were members of the Hong Kong Aircrew Officers Association (HKAOA) trade union.
- On 9^{th} July 2001 there were 1,555 pilots on the CPA seniority list.
- 1,300 of the 1,555 pilots were members of the HKAOA.
- The 2001 General Committee of the HKAOA comprised 20 pilots; 5 of them were terminated.

- The combined Year 2000 and 2001 General Committees of the HKAOA comprised of 31 pilots (several serving on both Committees); 9 of them were terminated.
- The HKAOA Negotiating Committee comprised 7 pilots ; 4 of them were terminated.

Issues in Dispute:

The 49ers claim that they were selected for dismissal on the basis of HKAOA trade union membership, HKAOA Committee involvement and/or participation in the HKAOA Negotiating Committee. At various times, in its defense, CPA denies such claim. This can be taken to imply that the 49ers were selected *randomly.*

Objective:

The aim of the analysis is to examine the probability of these events as if occurred randomly.

Data

We make use of the dismissal data supplied by the 49ers and the company data from Cathy Pacific.

Method

Hypergeometric Distribution:

The Hypergeometric Distribution (Annex I, Everitt, 1995) is a discrete probability distribution that describes the chances of selecting some individuals from a finite population without replacement. The method can be illustrated by calculating the probabilities of winning various prizes of the Mark Six Lottery where a sample of numbers (in this case up to 6) is selected without replacement from a greater population of numbers (in this case 49). For example, in order to win the jackpot any player need to select six numbers which are the same as the chosen six and the probability is (1 in 13.9 million, see Appendix II)

Here, we apply the Hypergeometric Distribution to find out the probability of the four events in the disputes as if occurred randomly:

1. What is the probability if randomly selected that 49 or more of the 50 pilots would be members of the HKAOA?

2. What is the probability if randomly selected that the 49ers would comprise 5 or more of the Year 2001 HKAOA General Committee?

3. What is the probability if randomly selected that the 49ers would comprise 4 or more of the HKAOA trade union Negotiating Committee?

4. What is the probability if randomly selected that the 49ers would comprise 9 or more of the combined Year 2000 and/or 2001 HKAOA General Committees?

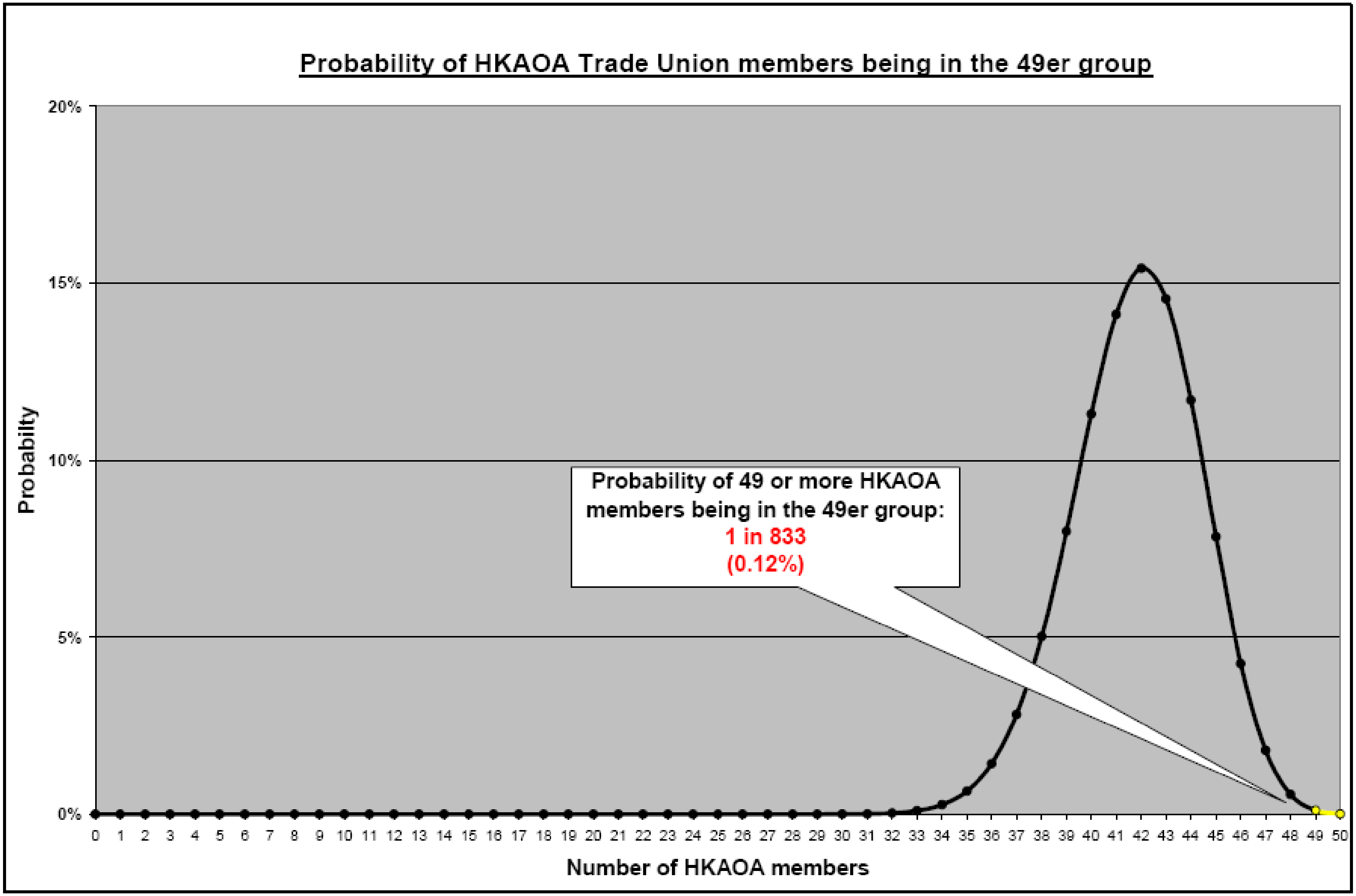
To facilitate comparison, the probabilities of selecting winning numbers from the 49 numbers in the Mark Six lottery are also calculated using the Hypergeometric Distribution.

Table 1: Probability of being in the 49er group* (decreasing likelihood)

49 or more HKAOA trade union members	**0.12%**	**1 in 833**
Choosing 4 winning Mark Six lottery numbers	**0.097%**	**1 in 1,031**
5 or more members of the Year 2001 General Committee	**0.030%**	**1 in 3,333**
4 or more members of the HKAOA Negotiating Team	**0.0031%**	**1 in 32,258**
Choosing 5 winning Mark Six lottery numbers	**0.0018%**	**1 in 55,556**
9 or more members of Year 2000 or 2001 General Committee	**0.000021%**	**1 in 4,761,905**
Choosing 6 winning Mark Six lottery numbers	**0.0000072%**	**1 in 13,983,816**

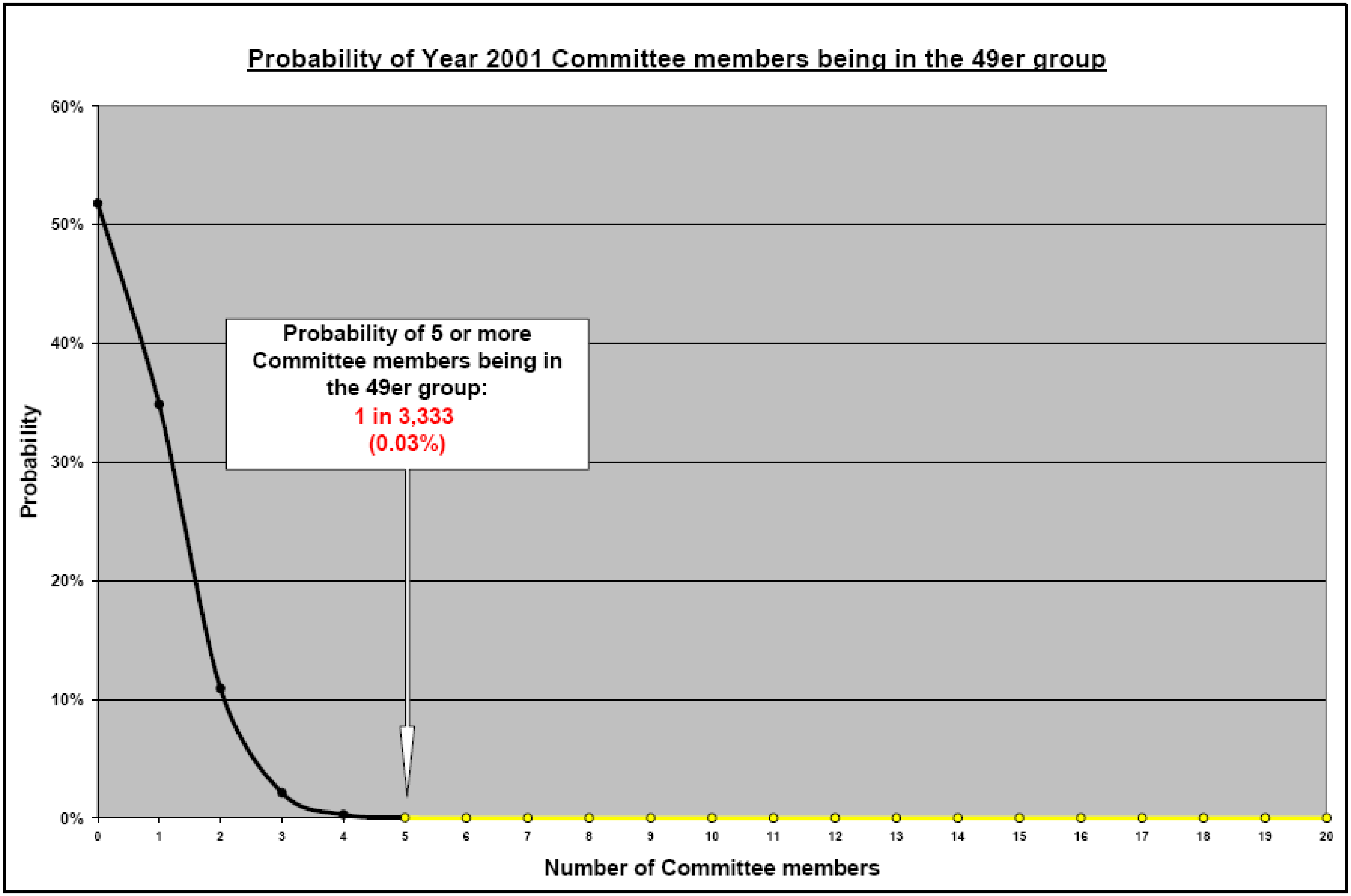
*see Appendix I and II for details of calculation

Scenario 1

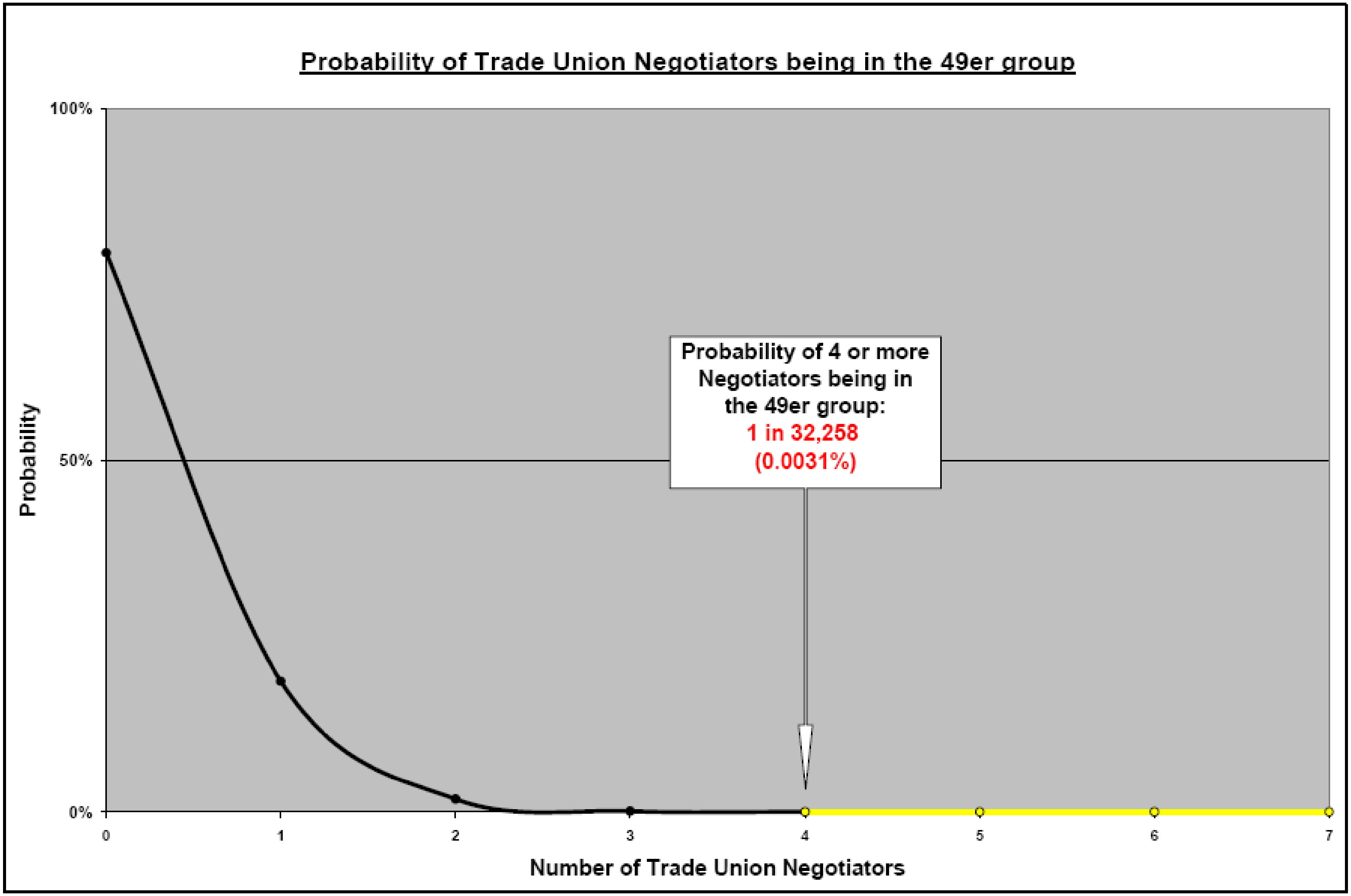


Probability of HKAOA Trade Union members being in the 49er group
Probability
20%
15%
10%
5%
0%
Probability of 49 or more HKAOA members being in the 49er group:
1 in 833
(0.12%)
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
Number of HKAOA members

Scenario 2

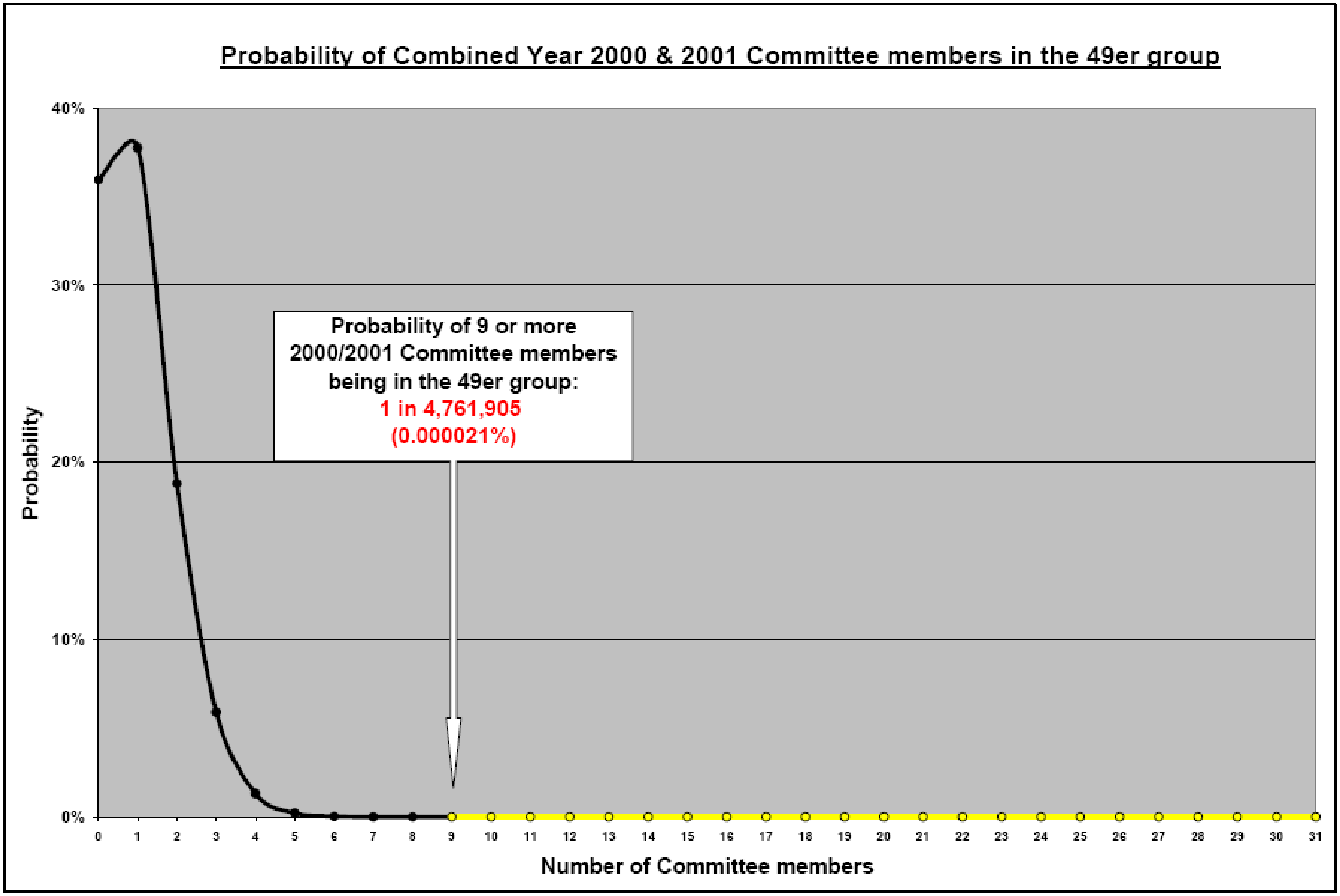

Probability of Year 2001 Committee members being in the 49er group
60%
50%
40%
30%
20%
10%
0%
Probability
Probability of 5 or more Committee members being in the 49er group:
1 in 3,333
(0.03%)
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
Number of Committee members

Scenario 3



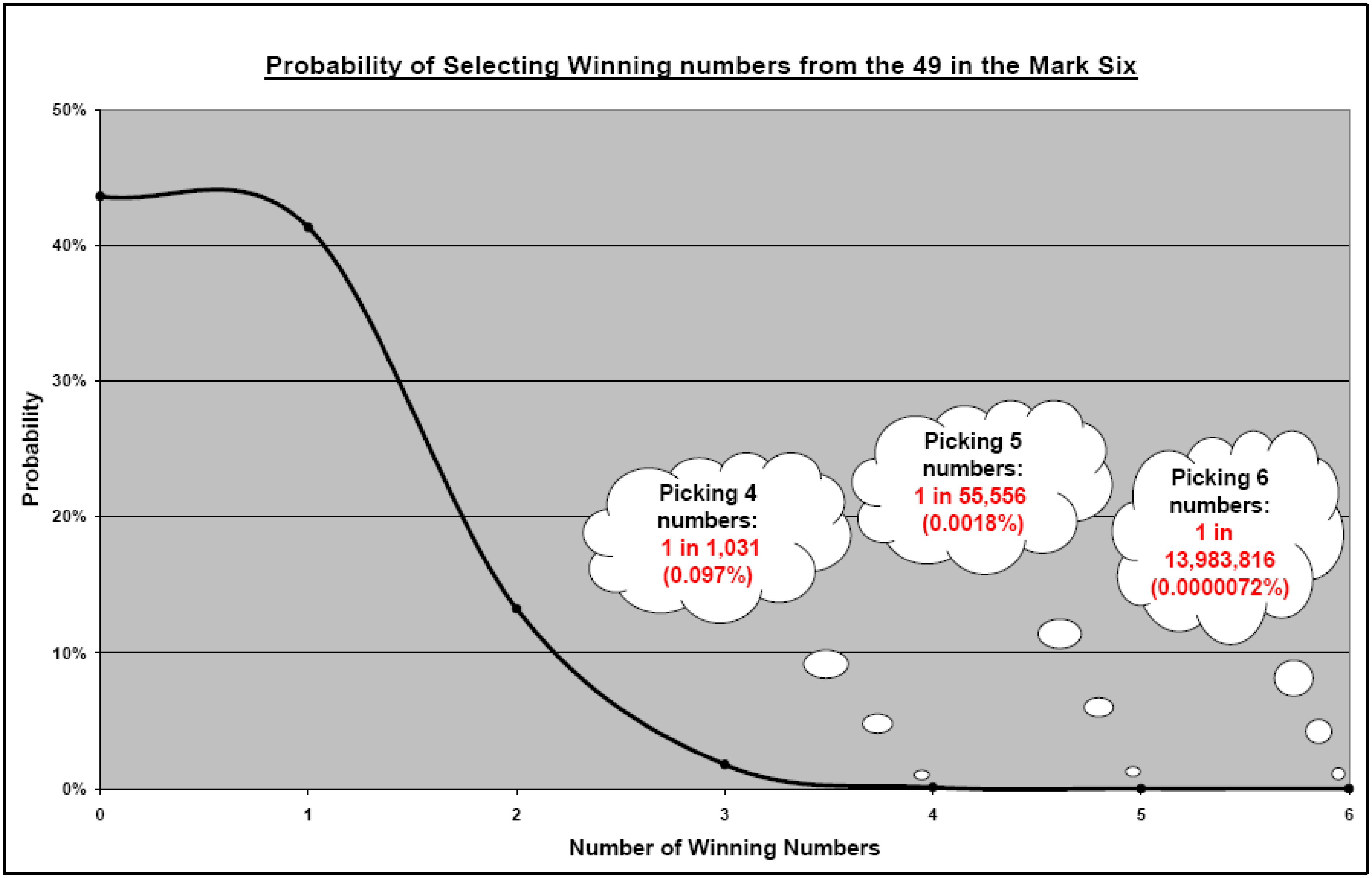
Probability of Trade Union Negotiators being in the 49er group
100%
50%
0%
Probability
0
1
2
3
4
5
6
7
Number of Trade Union Negotiators
Probability of 4 or more Negotiators being in the 49er group:
1 in 32,258
(0.0031%)

Scenario 4

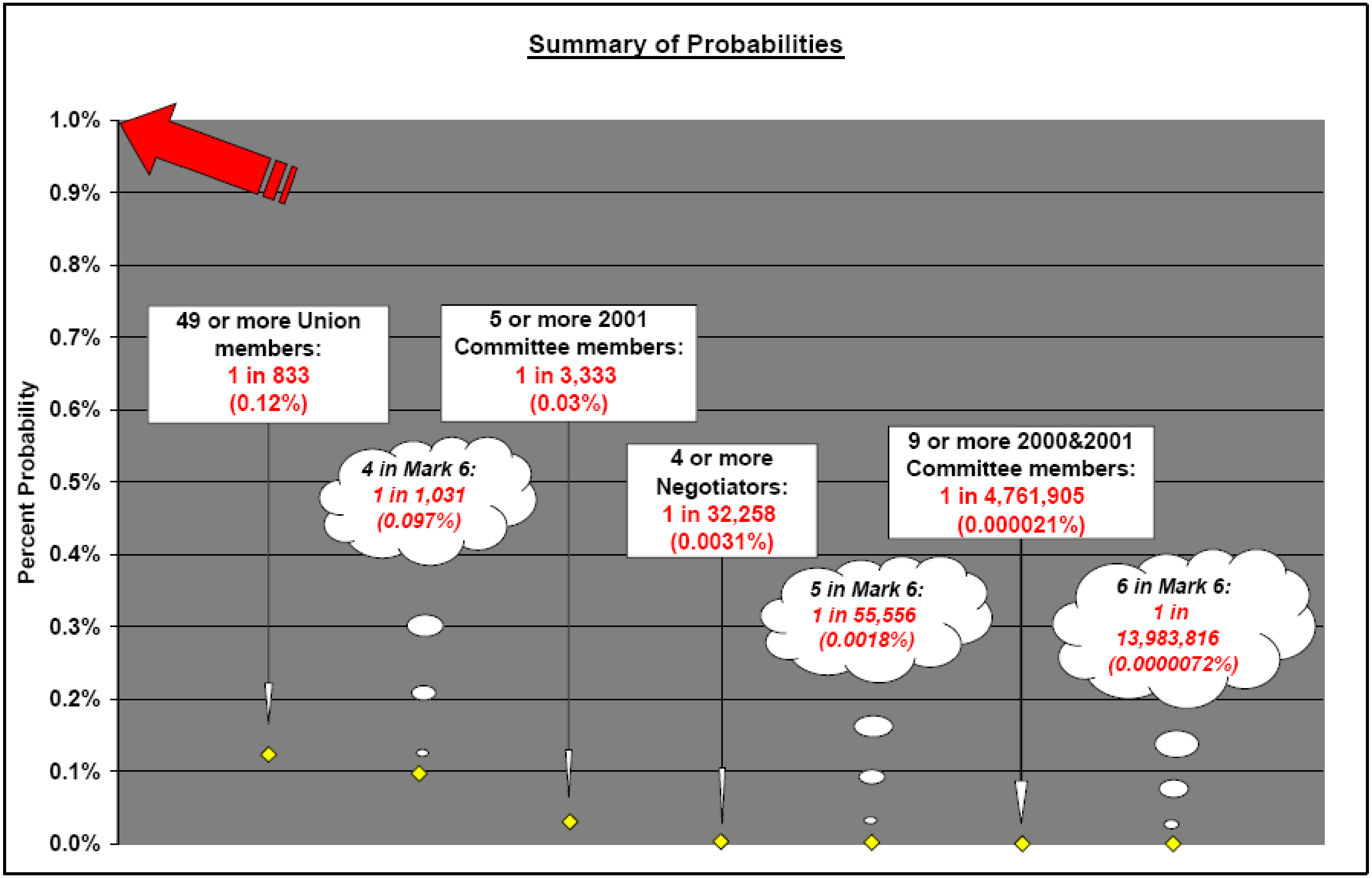


Probability of Combined Year 2000 & 2001 Committee members in the 49er group
Probability
40%
30%
20%
10%
0%
Probability of 9 or more 2000/2001 Committee members being in the 49er group:
1 in 4,761,905
(0.000021%)
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31
Number of Committee members

Mark Six Lottery Hypergeometric Distribution



Summary



Summary of Probabilities
Percent Probability
1.0%
0.9%
0.8%
0.7%
0.6%
0.5%
0.4%
0.3%
0.2%
0.1%
0.0%
49 or more Union members: 1 in 833 (0.12%)
4 in Mark 6: 1 in 1,031 (0.097%)
5 or more 2001 Committee members: 1 in 3,333 (0.03%)
4 or more Negotiators: 1 in 32,258 (0.0031%)
5 in Mark 6: 1 in 55,556 (0.0018%)
9 or more 2000&2001 Committee members: 1 in 4,761,905 (0.000021%)
6 in Mark 6: 1 in 13,983,816 (0.0000072%)

Results:

The probability that 49 or more of the "49ers" were HKAOA trade union members is 1 in 833. It is unlikely that this outcome is the result of random selection.

The probability that 5 or more of the 20 members of the Year 2001 HKAOA General Committee were in the 49er group is 1 in 3,333. It is very unlikely that this outcome is the result of random selection.

The probability that 4 or more of the 7 members of the HKAOA Negotiating Committee were in the 49er group is 1 in 32,258. This is a similar probability to that of selecting 5 winning numbers out of the 49 numbers in the Mark Six lottery. It is exceedingly unlikely that this outcome is the result of random selection.

The probability that 9 or more of the 31 members of the combined Year 2000 and 2001 HKAOA General Committees were in the 49er group is 1 in 4,761,905. It borders on the impossibility that this outcome is the result of random selection.

It is usual practice to set a cutoff value of 0.05 as the usual benchmark to reject the hypothesis of random dismissal (Everitt, 1995). And the associated p-value is calculated based on the observed value or more extreme than that to examine the claim of random dismissal.

All the resulting probabilities are much smaller than 0.05. Hence it is very unlikely that these events occurred randomly.

Conclusion

The suggestion that the four events were the result of random selection are very unlikely and the probability value is extremely small..

Annex I - Hypergeometric Distribution

[From Wikipedia, 2007]

In probability theory and statistics, the **hypergeometric distribution** is a discrete probability distribution that describes the number of successes in a sequence of *n* draws from a finite population without replacement.

	drawn	not drawn	total
defective	k	$D - k$	D
non-defective	$n - k$	$N + k - n - D$	$N - D$
total	n	$N - n$	N

A typical example is illustrated by the contingency table above: there is a shipment of *N* objects in which *D* are defective. The hypergeometric distribution describes the probability that in a sample of *n* distinctive objects drawn from the shipment exactly *k* objects are defective.

In general, if a random variable *X* follows the hypergeometric distribution with parameters *N*, *D* and *n*, then the probability of getting exactly *k* successes is given by

$$f(k;N,D,n) = \frac{\binom{D}{k}\binom{N-D}{n-k}}{\binom{N}{n}}$$

The probability is positive when k is between max{ 0, $D + n - N$ } and min{ n, D }.

The formula can be understood as follows: There are $\binom{N}{n}$ possible samples (without replacement). There are $\binom{D}{k}$ ways to obtain k defective objects and there are $\binom{N-D}{n-k}$ ways to fill out the rest of the sample with non-defective objects.

When the population size is large compared to the sample size (i.e., N is much larger than n) the hypergeometric distribution is approximated reasonably well by a binomial distribution with parameters n (number of trials) and $p = D / N$ (probability of success in a single trial).

Appendix II

The probability of winning various prizes of the Mark Six, where a sample of numbers (in this case up to 6) is selected without replacement from a greater population of numbers (in this case is 49), can be calculated according to the Hypergeometric Distribution as stated in Annex I. To be specific, the parameters for the Hypergeometric Distribution for the Mark Six are $N = 49$, $D = 6$, $n = 6$. Moreover, k is the number of winning numbers, for instance, $k = 6$ is for the first prize; $k = 5$ with a supplementary number is for second prize; and k=5 without the supplementary number is for the third prize and $k = 4$ for the fifth prize.

$$\Pr(\mathit{win\ the\ first\ prize}) = f(k = 6, N, D, n) = \frac{\binom{6}{6}\binom{43}{0}}{\binom{49}{6}} = \frac{1}{\binom{49}{6}} = 0.0000072\%$$

$$\Pr(\mathit{win\ the\ thrid\ prize}) = f(k = 5, N, D, n) = \frac{\binom{6}{5}\binom{43}{1}}{\binom{49}{6}} = 0.0018\%$$

$$\Pr(\mathit{win\ the\ fifth\ prize}) = f(k = 4, N, D, n) = \frac{\binom{6}{4}\binom{43}{2}}{\binom{49}{6}} = 0.097\%$$

Appendix III

1. Under Scenario 1 that the randomly selected 49 or more of the 50 pilots would be members of the HKAOA, N=1555, D=1300, n=50, k >= 49:

$$\Pr = f(k \geq 49, 1555, 1300, 50) = \sum_{i=49}^{50} \frac{\binom{1300}{i}\binom{255}{50-i}}{\binom{1555}{50}} = 0.12\%$$

2. Under scenario 2 that the randomly selected 49ers would comprise 5 or more of the Year 2001 HKAOA General Committee, N=1555, D=20, n=50, k >= 5:

$$\Pr = f(k \geq 5, 1555, 20, 50) = \sum_{i=5}^{20} \frac{\binom{1535}{50-i}\binom{20}{i}}{\binom{1555}{50}} = 0.030\%$$

3. Under scenario 3 that the randomly selected 49ers would comprise 4 or more of the HKAOA trade union Negotiating Committee, N=1555, D=7, n=50, k >= 4:

$$\Pr = f(k \geq 4, 1555, 7, 50) = \sum_{i=4}^{7} \frac{\binom{1548}{50-i}\binom{7}{i}}{\binom{1555}{50}} = 0.0031\%$$

4. Under scenario 4 that the randomly selected 49ers would comprise 9 or more of the combined Year 2000 and/or 2001 HKAOA General Committees, N=1555, D=31, n=50, k >= 9:

$$\Pr = f(k \geq 9, 1555, 31, 50) = \sum_{i=9}^{31} \frac{\binom{1524}{50-i}\binom{31}{i}}{\binom{1555}{50}} = 0.000021\%$$

References

[1] Everitt, B.S. (1995), *The Cambridge Dictionary of Statistics in the Medical Sciences*, 1st ed., Cambridge, England: Cambridge University Press.

[2] Wikipidia, 2007.