The Ideal Number of Questions to Attempt Randomly for Maximum Positive Marks

Author: *Kaju Burfi*

Consider you're writing an exam. It has only Multiple Choice (Single correct) Questions [MCQs]. What would be the best number of questions to *guess*? Obviously, this would depend upon the marking scheme.

If there were no negative marking, i.e. even if you get an answer wrong, you get zero marks. In this case, the best bet would be to attempt all questions.

If there **is** negative marking, things get complicated. Consider a common marking scheme:

- Correct answer: **+4**
- Wrong answer: **−1**
- Unanswered: $\pm$**0**

The questions each have 4 options, only one of them are correct. Hence, if we were to guess randomly, we have a 1 in 4 chance of getting it right(**+4**), or a 3 in 4 chance of getting it wrong(**−1**). Let n be the number of questions we *guess*.

We can calculate the expected value, μ, as the product of the probability of getting a particular mark and the value of the mark we attain.

For $n = 1$, we have two possible cases(as discussed above). Hence, the expected value,

$$\mu_1 = \left(-1 \times \frac{3}{4}\right) + \left(4 \times \frac{1}{4}\right) = \frac{1}{4} \tag{1}$$

For $n = 2$, we have three possible cases:

- Both wrong ⟶ (-1) + (-1) = −2 — Probability = $\frac{3}{4} \times \frac{3}{4} = \frac{9}{16}$
- One correct, one wrong ⟶ (-1) + (4) = 3 — Probability = $\frac{3}{4} \times \frac{1}{4} = \frac{6}{16}$
- Both correct ⟶ (4) + (4) = 8 — Probability = $\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$

Calculating the expected value, we get:

$$\mu_2 = \frac{1}{2} \tag{2}$$

This means that on average, if we were to *guess* two questions, we would get half a mark. But we cannot get **half a mark** in an actual exam(of this format).

Table 1 gives the probabilities for some values of n:

<table>
<tr><th>n</th><th colspan="4">Possible marks & their probabilities</th><th>Expected value(μ)</th></tr>
<tr><td rowspan="2">1</td><td colspan="2">−1</td><td colspan="2">4</td><td rowspan="2">$\frac{1}{4}$</td></tr>
<tr><td colspan="2">(3/4)</td><td colspan="2">(1/4)</td></tr>
<tr><td rowspan="2">2</td><td>−2</td><td>3</td><td colspan="2">8</td><td rowspan="2">$\frac{1}{2}$</td></tr>
<tr><td>(9/16)</td><td>(6/16)</td><td colspan="2">(1/16)</td></tr>
<tr><td rowspan="2">3</td><td>−3</td><td>2</td><td>7</td><td>12</td><td rowspan="2">$\frac{3}{4}$</td></tr>
<tr><td>$(3/4)^3$</td><td>$3(3/4)^2(1/4)$</td><td>$3(3/4)(1/4)^2$</td><td>$(1/4)^3$</td></tr>
<tr><td rowspan="2">n</td><td>$-n$</td><td>$4-(n-1)$</td><td>$8-(n-2)$ ...</td><td>... $4n$</td><td rowspan="2">$\frac{n}{4}$</td></tr>
<tr><td>$\binom{n}{0}\left(\frac{3}{4}\right)^n$</td><td>$\binom{n}{1}\left(\frac{3}{4}\right)^{n-1}\left(\frac{1}{4}\right)$</td><td>$\binom{n}{2}\left(\frac{3}{4}\right)^{n-2}\left(\frac{1}{4}\right)^2$</td><td>$\binom{n}{n}\left(\frac{1}{4}\right)^n$</td></tr>
</table>

Table 1: Possible marks and their probabilities

Let us denote the marks obtained, given we got i correct, as x_i. Hence, if we get zero correct(or all wrong), we get $x_0 = -n$ marks, where n is the number of questions attempted.

In general,

$$x_i = 4r - (n - r) = 5r - n \qquad \text{for } 0 \le r \le n \tag{3}$$

Here, r is a dummy variable which we iterate over to get the total probability.

We would like to get non-negative marks, or

$$x_i \geq 0 \tag{4(1)}$$

$$\iff r \geq \frac{n}{5} \tag{4(2)}$$

But,

$$r, n \in \mathbb{Z} \Rightarrow r \geq \left\lceil \frac{n}{5} \right\rceil \tag{5}$$

Where $\lceil \cdot \rceil$ is the ceiling function.

We want the probability of getting non-negative marks, i.e.

$$\mathrm{P}(x_i \geq 0) = \sum_{r=\lceil n/5 \rceil}^{n} \binom{n}{r} \left(\frac{3}{4}\right)^{n-r} \left(\frac{1}{4}\right)^{r} \tag{6(1)}$$

$$= \left(\frac{3}{4}\right)^{n} \sum_{r=\lceil n/5 \rceil}^{n} \frac{\binom{n}{r}}{3^r} \tag{6(2)}$$

This is just a bernoulli's trials, with $p = \frac{1}{4}$ and $q = \frac{3}{4}$.

Computing this for various values of n, we get the following:

n	$P(x_i \geq 0)$
1	25%
2	43.75%
3	57.81%
4	68.35%
5	76.27%
6	46.60%
7	55.50%
8	63.20%
9	69.96%
10	75.59%
11	54.47%

Table 2: Probability of getting non-negative marks

Here we see that the best possible number of questions to attempt would be a multiple of 5. But this would be quite misleading. It is true, but misleading.

We considered $P(x_i \geq 0)$, which includes the cases where $x_i = 0$. This case occurs only when $n = 5k, k \in \mathbb{N}$. In fact, when you calculate the probabilities, $P(x_i = 0)$ is very high. Some values are given in Table 3:

n	$P(x_i = 0)$
5	39.55%
10	28.15%
15	22.51%

Table 3: Probability of getting zero marks

Now, considering this as well, we can make the final table to get the probabilities of getting *positive* marks.

n	$P(x_i > 0)$
1	25%
2	43.75%
3	57.81%
4	68.35%
5	36.72%
6	46.60%
7	55.50%
8	63.20%
9	69.96%
10	47.44%
11	54.47%

Table 4: Probability of getting positive marks

Some observations from Table 4:

- Probability of getting a zero is more than that of getting positive marks if $n = 5$.

$$\mathrm{P}(x_i = 0) > \mathrm{P}(x_i > 0) \text{ when } n = 5 \quad (7(1)$$
$$[39.55\%] \qquad [36.72\%] \quad (7(2)$$

- Best number of questions n to attempt is $5k - 1, k \in \mathbb{N}$, i.e. 4, 9, 14, etc.

Conclusion

For an exam with a marking scheme as +4 and −1, the best possible number of questions to answer would be a one less than a multiple of 5, or $5k - 1, k \in \mathbb{N}$, with probabilities of getting postive marks as 68.35% for 4 questions and 69.96% for 9 questions.

Note that this assumes completely random choice of options. In reality, you will generally have some idea, which reduces the probability of negatives. So, it is almost always better than 70% chance, given you attempt 4 or 9 questions.

So, take ~~risks~~ calculated risks!