A Preliminary Study of Gender Differences in Participation Rates in Competitive Programming Contests

Vishakha Agrawal, Simran Bhurat
Department of ISE
Dayananda Sagar College of Engineering, Bengaluru
vishakha.agrawal09@gmail.com, bhuratsimran@gmail.com

Oindrilla Banerjee, Vindhya Malagi
Department of CSE
Dayananda Sagar College of Engineering, Bengaluru
ob.oindrila.banerjee@gmail.com, vindhya-cs@dayanandasagar.edu

Viraj Kumar
Department of CSE
Dayananda Sagar University
Bengaluru
viraj-cse@dsu.edu

1 INTRODUCTION

In 2009, an influential survey of Indian IT businesses conducted by NASSCOM and McKinsey found that a mere 26% of Indian engineering graduates were employable [1]. A more recent joint study by Aspiring Minds and NASSCOM found that employability remains low [2]. Aspiring Minds specializes in filtering job applicants, and it has broken down employability data from 2016 [3] and 2019 [4] according to multiple factors, including gender. Both these reports find that females are less employable than males (as per Aspiring Minds' objective testing criteria), and the percentage of employable females appears to have declined from 2016 to 2019 in both product-based companies (from 2.71% to 2.54%) and service-based companies (from 17.5% to 15.5%). Note that this issue is distinct from the well-studied "pipeline shrinkage problem", which is concerned with the decline in females (relative to males) entering computing programmes [7].

We find this data troubling, particularly because we are unaware of such trends based on academic evaluation criteria. In our own computing departments (Information Science and Engineering, and Computer Science and Engineering), for instance, female students are well represented among the best-performing students (in terms of GPA), even though they are under-represented in terms of enrollment. Thus, we believe that an important question for CS education research in India is: What explains the apparent discrepancy between female students' on-par (or better) academic performance relative to male students and their sub-par employability? Such explanations may inform strategies to improve the employability of female graduates in computing programmes.

The NASSCOM-Aspiring Minds study [2] identifies several reasons for low employability, including a "lack of basic coding skills" among job applicants. A separate Aspiring Minds report [5] finds that males outperform females on programming tasks not only in premier (Tier-1) institutions, but in institutions such as ours that lie outside this top tier as per the NIRF rankings [6]. Again, we are unaware of a consistent gap reported between females and males on exams, quizzes, and assignments in programming-intensive courses, including within the departments represented by the authors.

It is possible that there is a mismatch between industry needs (as reflected in assessments by companies such as Aspiring Minds) and in-course programming assignments. Mindful of this, and to provide students additional opportunities to hone their technical skills primarily in programming, data structures, and algorithms (all of which are in high demand by employers [8]), many institutions organize hackathons and competitive programming contests for students. There is anecdotal evidence that female participation in such contests is lower than male participation, even after adjusting for differences in enrollment rates. If so, this may provide a partial explanation for the gender differences in programming abilities [5] and employability in the IT industry [2].

The present study examines participation data from a series of competitive programming contests, *PB Hustles*, organized by Point Blank, a community established by the Computer Science and Engineering department of Dayananda Sagar College of Engineering (DSCE).

2 METHODOLOGY

A. Data source. Our data has been collected from *PB Hustles* contests over a period of six months. The contests are conducted on Code Forces (https://codeforces.com), a popular online competitive programming platform. These contests are open only to DSCE students. Although we see some participation from Electronics and Communication Engineering and Mechanical Engineering students, most of the participants are from the departments of Computer Science and Engineering (CSE) and Information Science and Engineering (ISE). We present an analysis of data only from undergraduate students in these departments. We further restrict our analysis by eliminating contests where the total number of participants was below 10. This leaves us with data from 13 different contests, as shown in Table 1.

Table 1. Participation details of female and male students from ISE and CSE for 13 PB Hustle contests

PB Hustle Contest No.	ISE participants		CSE participants	
	Female	Male	Female	Male
2.4	0	10	5	18
2.5	1	6	2	15
2.6	1	7	2	12
2.7	1	5	0	14
2.8	1	4	2	8
2.10	0	6	0	11
2.12	0	4	3	10
2.15	2	6	2	7
2.18	0	5	2	4
2.19	1	5	1	4
2.20	0	4	2	5
2.21	0	4	2	6
2.22	0	3	1	7

B. Metrics. A cursory look at our data reveals that the number of female participants is substantially lower than the number of male participants. However, there are fewer female students than male students enrolled in both the ISE department (128 females vs 266 males) and the CSE department (178 females vs 555 males). Thus, for each contest *c*, department $d \in$ {ISE, CSE} and gender $g \in$ {female, male}, we define the normalized participation rate $p(c, d, g)$ as the percentage of the total number of students of gender *g* across *all* semesters in department *d* who participated in the contest *c*. For our analysis, we compare the average value of $p(c, d, g)$ across all 13 contests:

$$F_d \stackrel{\text{def}}{=} avg_c\ p(c, d, \text{female})$$
$$M_d \stackrel{\text{def}}{=} avg_c\ p(c, d, \text{male})$$

Our null hypotheses are:

- $H_{ISE}: F_{ISE} = M_{ISE}$ (1)
- $H_{CSE}: F_{CSE} = M_{CSE}$ (2)

3 ANALYSIS

In our data, all females from ISE were from a single semester (4th semester), whereas males from ISE and CSE as well as females from CSE were from the 2nd, 4th, and 6th semesters. Since the contests were open to *all* students, we have computed the percentages $p(c, d, g)$ exactly as defined above, where the denominator is the total number of students of gender *g* across all semesters of department *d*.

The contests are numbered chronologically, and we observe a steady decline in the total number of participants across all contests. (As noted earlier, contests where the total number of participants is below 10 are discarded altogether.)

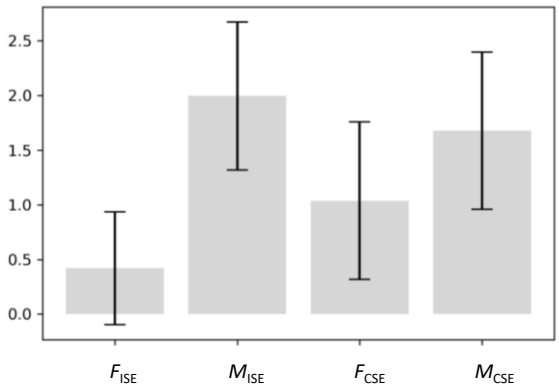

Figure 1. The metrics F_{ISE}, M_{ISE}, F_{CSE}, and M_{CSE} with 95% confidence intervals.

It is clear from Figure 1 that in the ISE department, females participate at substantially lower rates (just over 0.4% on average) than males (nearly 2% on average). Bearing in mind that our data is limited, a t-test for the null hypothesis H_{ISE} (1) results in a *p*-value less than 0.0001 (extremely significant), suggesting that this hypothesis can be rejected. In the CSE department, although females once again participate at lower rates (just over 1% on average) than males (almost 1.7% on average), the t-test for the null hypothesis H_{CSE} (2) results in a *p*-value of 0.0357. Hence, it is less clear whether this null hypothesis can be rejected.

4 DISCUSSION AND CONCLUSIONS

We acknowledge a key limitation of our data: it has been gathered from only a single institution (DSCE), although we have managed to obtain data from two different computing departments within this institution (ISE and CSE). Hence, we can only draw tentative conclusions based on this data.

Our results are in line with anecdotal evidence that female participation rate in competitive programming contests is lower than male participation rate. Our analysis technique based on normalized participation rates accounts for differences in enrollment rates by gender, and it should be applicable at all institutions. Indeed, we would like to contribute towards a multi-institutional study of participation rates by gender so that our research community can gain deeper insight into this issue.

The gap between female and male participation rates is starker for ISE than CSE. We did consider whether this could be because *PB Hustle* contests, which are organized by the CSE department, attract fewer ISE students. However, we observe from Figure 1 that there are substantial overlaps between the 95% confidence intervals for F_{ISE} and F_{CSE}, as well as for M_{ISE} and M_{CSE} (in fact, $M_{ISE} > M_{CSE}$).

We note that the participation rate we have observed in competitive programming contests is very poor (less than 2%) for all groups. If this rate is similarly poor across many institutions, then even a substantial gap between

female and male participation rates cannot explain the difference in female and male programming ability as reported by Aspiring Minds [5]. Thus, our preliminary study has identified an interesting topic for subsequent research. Further, as we strive to improve participation rates in competitive programming contests, we will be able to investigate additional questions for which we do not have adequate data. One such question is whether the observed differences in participation rates by gender vary across semesters i.e., does the gender gap widen or shrink as students spend more time within our institution? Such analyses can help departments understand whether they are successful in providing learning opportunities for all students in an inclusive manner.

REFERENCES

[1] NASSCOM and McKinsey. 2009. Perspective 2020: Transform Business, Transform India.

[2] NASSCOM and Aspiring Minds. 2020. How IT Services Industry Hires: Skills & Competencies that are Imperative. https://www.aspiringminds.com/research-reports/how-IT-services-industry-hires-skills-and-competencies-that-are-imperative

[3] Aspiring Minds. 2016. National Employability Report for Engineers. https://www.aspiringminds.com/research-reports/national-employability-report-for-engineers-2016

[4] Aspiring Minds. 2019. National Employability Report for Engineers. https://www.aspiringminds.com/research-reports/national-employability-report-for-engineers-2019

[5] Aspiring Minds. 2017. National Programming Skills Report. https://www.aspiringminds.com/research-reports/national-programming-skills-report-for-engineers-2017

[6] National Institutional Ranking Framework, (2020), India Rankings 2020: Engineering https://www.nirfindia.org/2020/EngineeringRanking.html

[7] Denise W Guren, Tracy Camp, (2002), An ACM-W literature review on women in computing, ACM SIGCSE Bulletin 34(2):121-127.

[8] Prashant R Nair, (2020), Increasing Employability of Indian Engineering Graduates through Experiential Learning Programs and Competitive Programming: Case Study. In Elsevier Procedia Computer Science 172, 831–837.