Trend analysis of diversity in conferences

Ruijia Yang

15/07/2020

Introduction

This analysis is to identify whether artificial intelligence field are open to minorities of various sorts (gender, ethnic background) by mining related conferences, to see who is attending and explore how diverse these conferences are according to three factors: population, race and gender.

Methods

For gender diversity research, the analysis uses two-proportions z-test to compare two observed proportions, which are in the first year and last year of the conference.
For race diversity research, the analysis uses Chi-Square test of independence, assuming 0.1 significance level, in order to test if there is a difference between race distribution and years of different conference, What's more, the analysis conduct Chi-Square goodness of fit test on each conference, to compare multiple observed proportions to expected probabilities. Assume if these races were equally distributed, the expected proportion would be 1/4 for each of the race.

Data

All participants' names in this analysis are from AAAI Conference websites and Applied AI Conference websites by web scraping, and then the analysis use NamSor API v2 to analyze participants' gender (male or female) and likely race/ethnicity (W_NL (white, non latino), HL (hispano latino), A (asian, non latino), B_NL (black, non latino)), according to US Census taxonomy. In AAAI Conference, I categorize the areas where participants from into North America (including USA and Canada), Europe (including Germany, Greece, Italy, UK, Spain, Belgium, Netherlands, Sweden, Ireland, France, Portugal), Asia (including China, Japan, Korea and India) and Oceania. Due to the loss of data of all other participants except chairs and invited speakers in 2018, this analysis considers the incomplete data in 2018 as an outlier and exclude it from the analysis. Below showing the composition of different race in two conferences.

year	A	B_NL	HL	W_NL	total
2013	210	63	111	333	717
2014	162	39	74	258	533
2015	366	84	200	470	1120
2016	516	74	158	446	1194
2017	866	96	218	535	1715
2019	2382	114	327	630	3453
2020	4907	645	451	575	6578

Figure 1:AAAI Conference ethnicity composition

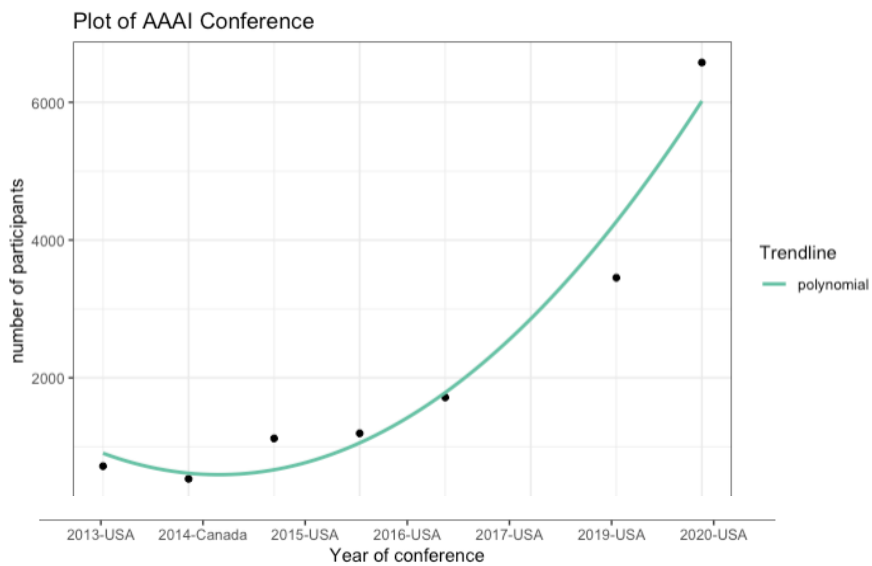
year	A	B_NL	HL	W_NL	total
2018	22	3	1	19	45
2019	34	2	5	23	64
2020	36	1	8	28	73
2021	9	1	3	4	17

Figure 2: Applied AI Conference ethnicity composition

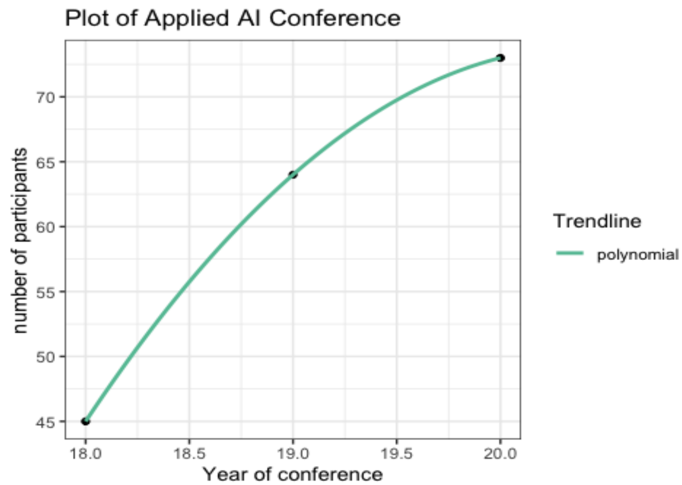
Analysis

1.Population

In terms of population diversity, For AAAI conference, due to the loss of data of all other participants except chairs and invited speakers in 2018, considering the incomplete data in 2018 as an outlier and exclude it from the analysis. There is an increasing trend of attendees in the AAAI conference from 2013 to 2020. AAAI, as a long-history academia conference, although the conferences mainly held in USA, there are a large group of committees every year from different places. Among participants in AAAI, there are only about 2% participants from Asia in 2013, while in 2020 there is at least 18% participants from Asia, and the other group of participants, for example, participants from North America, Europe and Oceania keeps constant. According to research, AAAI is holding more events to increase public understanding of artificial intelligence, improve the teaching and training of AI practitioners throughout these years, for example, more workshops in 2017 and 2019, adding a ‘Undergraduate Consortium’ program in 2020, which attracted more people to get involved in the conference.



For Applied AI conference, a newly industry conference, the number of participants is increasing through 2018 to 2020. Since it is more likely that an industry conference shows regional concentration trend, the conference is holding at 1 city in 2018, 6 cities in 2020, and going to be 11 cities in 2021. With the increasing number of participated cities, the degree of diversity of this conference can be proved.



Overall, the conferences in artificial intelligence field are having increasing attendees through these years, and more and more companies are participating, not only from different areas but different universities and companies. Therefore, taking the increasing population into account, conferences in artificial intelligence field are becoming more diverse through these years.

2.Gender

Talking about gender diversity, from the two-proportion test result for proportion of female in Applied AI conference in 2018 and 2020. Conducting power analysis on the two sample, using significance level 0.66 can be enough to detect some difference, where the confidence level is a little low, so the analysis stick with 0.1 significance level α. Since p-value > α, there is not enough evidence to reject null hypothesis that the proportion of females are not statistically different from 2018 to 2020, which suggests that the proportion of female in the industry conference stays unchanged.

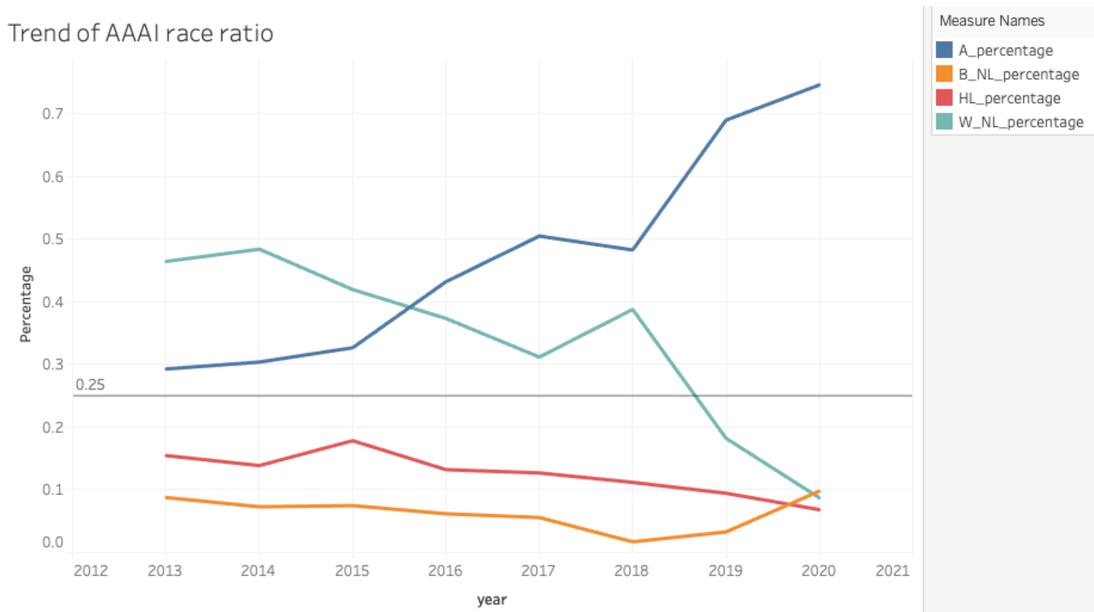
From the two-proportion test result for proportion of female in AAAI conference in 2013 and 2020, with 0.1 significance level α. Since p-value < α, there is enough evidance to reject null hypothesis that the proportion of female has not changed from 2013 to 2020, which suggests that the proportion of female in this conference is statistically different from 2013 to 2020, and the proportion of female is increasing through the years. And I am 90% confident that the true value of the difference between two proportion is in between 0.031 and 1.

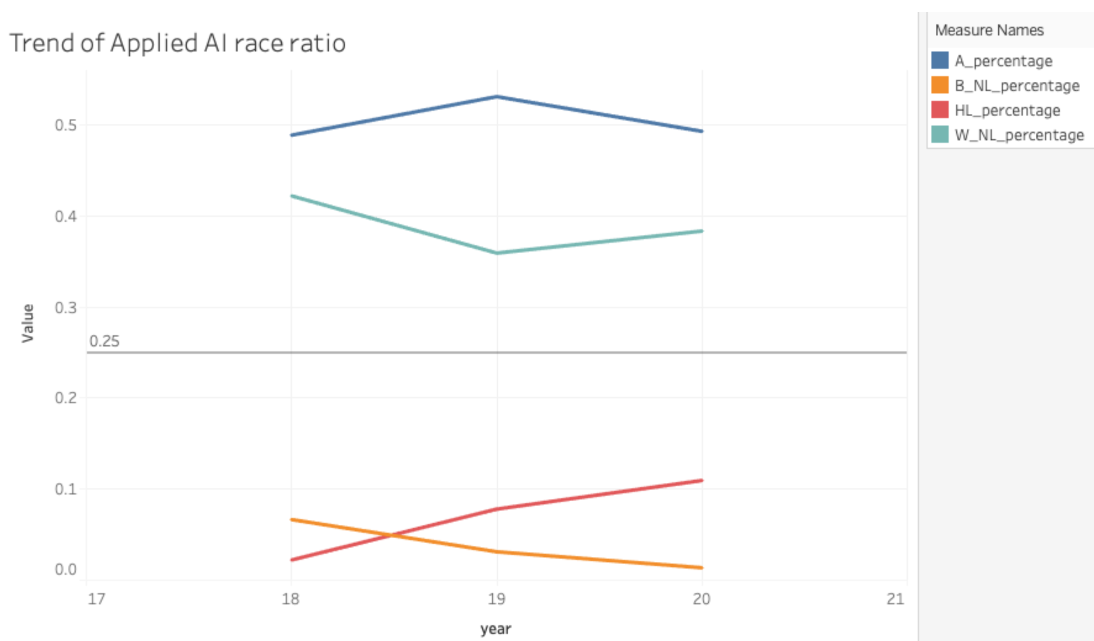
From the two-proportion test result for proportion of female for same year comparison on two conferences in 2018, 2019 and 2020, with a 0.1 confidence level α. Comparing conference in the same 3 years, the proportion of female in AAAI conference is always smaller than the proportion of female in Applied AI conference in 2018 and 2019. However, in 2020, p-value is larger than α, which means there is not a statistically difference between the proportion of female of two conferences in 2020.

Since gender diversity is equitable or fair representation of people of different genders. It most commonly refers to an equitable ratio of men and women, the result of the same year comparison tells that the proportion of female is increasing these years, who contributes to the gender diversity of conference in artificial intelligence field.

3.Race

Below showing the trend line of different race ratio in different conference. For AAAI conference, the number of Asian is increasing dramatically from 2013 to 2020, and the number of black, non latino also showing an increasing trend. While the percentage of White, non latino is decreasing significantly, the percentage of hispano latino is slightly decreasing. Adding a 0.25 reference line, assume that if each line approach 0.25 more, the participant composition of race for the conference is more equal. The ratio of Asian and White nearly composite the main part of participants (each above 0.25) For Applied AI conference, the number of hispano in increasing while the black, non hispano in decreasing and Asian and White, non latino dominate the participants through 2018 to 2020. We can see that there might be some issues in the race distribution of the two conference through these years.





From the test result of Chi-Square test of independence, assuming 0.1 significance level α. Because of p-value < α, there is significant evidence to reject null hypothesis that participants in each year and race are independent, which means race distribution and conference in different year are statistically significantly associated.

In order to know whether these races are equally distributed or not, using Chi-Square goodness of fit test. The test result p-value for both tests are less than the 0.1 significance level α, which means there is significant evidence to reject null hypothesis that there is no significant difference between the observed and the expected value (1/4). We can conclude that the races are significantly not commonly distributed and there is a significant difference between the observed and the expected value.

However, some people think racial diversity means different groups occupying the same physical space and is inter-twined with some form of racial inequality. The proportion of different race group is changing through these years, which implies that each race group has a chance to participate in these conferences and this can be seen as diversity in race factor of these conferences.

Overall, in terms of diversity of conference in the artificial intelligence field, the population and gender part are becoming more diverse. While there is not absolute equality in race distribution through these conferences from 2013 to 2020, the domination of participant group is changing, which also shows a trend of race diversity.

Appendix

```
library("readxl")
aoolied_file <- read_excel("/Users/ricayang/Desktop/applied_race.xlsx")
applied_ai_x <- aoolied_file$year
applied_ai_y <- aoolied_file$total
library("ggplot2")
DF <- data.frame(applied_ai_x, applied_ai_y)

ggplot(DF, aes(x = applied_ai_x, y = applied_ai_y)) +
  geom_point() + ggtitle("Plot of Applied AI Conference") +
  scale_x_continuous(name="Year of conference") +
  scale_y_continuous(name="number of participants")+
  stat_smooth(method = 'lm', formula = y ~ poly(x,2), aes(colour = 'polynomial'), se= FALSE)  + theme_b
w() +
  scale_colour_brewer(name = 'Trendline', palette = 'Set2')

aaai_file <- read_excel("/Users/ricayang/Desktop/aaai_race.xlsx")
aaai_x <- aaai_file$year
aaai_y <- aaai_file$total

DF2 <- data.frame(aaai_x, aaai_y)
ggplot(DF2, aes(x = aaai_x, y = aaai_y)) +
  geom_point() + ggtitle("Plot of AAAI Conference") +
  scale_x_continuous(name="Year of conference", limits=c(2013, 2020)) +
  scale_y_continuous(name="number of participants")+
  stat_smooth(method = 'lm', formula = y ~ poly(x,2), aes(colour = 'polynomial'), se= FALSE)  + theme_b
w() +
  scale_colour_brewer(name = 'Trendline', palette = 'Set2')

pwr.2p2n.test(h = 0.2, n1 = 45, n2 = 74, sig.level = NULL, power = 0.8)
##      difference of proportion power calculation for binomial distribution (arcsine transformation)
##
##              h = 0.2
##             n1 = 45
##             n2 = 74
##      sig.level = 0.6618683
##          power = 0.8
##    alternative = two.sided
##
## NOTE: different sample sizes

x <- factor(X18_file$gender,levels=c("male","female"), labels=c(0,1))
y <- factor(X20_file$gender,levels=c("male","female"), labels=c(0,1))
prop.test(x = c(sum(x == 1),sum(y == 1)), n = c(45,74), alternative = "greater",conf.level = 0.34)

##
##  2-sample test for equality of proportions with continuity correction
##
## data:  c(sum(x == 1), sum(y == 1)) out of c(45, 74)
## X-squared = 0.88282, df = 1, p-value = 0.1737
## alternative hypothesis: greater
```

```
## 34 percent confidence interval:
##  0.1170556 1.0000000
## sample estimates:
##    prop 1    prop 2
## 0.3555556 0.2567568
```

```
x <- factor(X13_aaai_file$gender,levels=c("male","female"), labels=c(0,1))
y <- factor(X20_aaai_file$gender,levels=c("male","female"), labels=c(0,1))
aaai_trend <- prop.test(x = c(sum(x == 1),sum(y == 1)), n = c(720,6581), alternative = "less",conf.level =
0.9)
##  2-sample test for equality of proportions with continuity correction
##
## data:  c(sum(x == 1), sum(y == 1)) out of c(720, 6581)
## X-squared = 9.6894, df = 1, p-value = 0.0009267
## alternative hypothesis: less
## 90 percent confidence interval:
##  -1.00000000 -0.03191394
## sample estimates:
##    prop 1    prop 2
## 0.1861111 0.2385656
```

same year comparsion

```
x <- factor(X18_aaai_file$gender,levels=c("male","female"), labels=c(0,1))
y <- factor(X18_file$gender,levels=c("male","female"), labels=c(0,1))
year18_trend <- prop.test(x = c(sum(x == 1),sum(y == 1)), n = c(113,45), alternative = "less",conf.level =
 0.9)
year18_trend
```

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  c(sum(x == 1), sum(y == 1)) out of c(113, 45)
## X-squared = 1.9929, df = 1, p-value = 0.07902
## alternative hypothesis: less
## 90 percent confidence interval:
##  -1.000000000 -0.005348504
## sample estimates:
##    prop 1    prop 2
## 0.2300885 0.3555556
```

```
x <- factor(X19_aaai_file$gender,levels=c("male","female"), labels=c(0,1))
y <- factor(X19_file$gender,levels=c("male","female"), labels=c(0,1))
year19_trend <- prop.test(x = c(sum(x == 1),sum(y == 1)), n = c(3457,64), alternative = "less", conf.level
 = 0.9)
year19_trend
```

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  c(sum(x == 1), sum(y == 1)) out of c(3457, 64)
```

```
## X-squared = 2.7329, df = 1, p-value = 0.04915
## alternative hypothesis: less
## 90 percent confidence interval:
##  -1.00000000 -0.01419927
## sample estimates:
##    prop 1    prop 2
## 0.2597628 0.3593750
```

```
x <- factor(X20_aaai_file$gender,levels=c("male","female"), labels=c(0,1))
y <- factor(X20_file$gender,levels=c("male","female"), labels=c(0,1))
year20_trend <- prop.test(x = c(sum(x == 1),sum(y == 1)), n = c(6581,74), alternative = "less", conf.level
 = 0.9)
year20_trend
```

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  c(sum(x == 1), sum(y == 1)) out of c(6581, 74)
## X-squared = 0.051942, df = 1, p-value = 0.4099
## alternative hypothesis: less
## 90 percent confidence interval:
##  -1.00000000  0.05406878
## sample estimates:
##    prop 1    prop 2
## 0.2385656 0.2567568
```

```
# aaai_race_dif
tulip2 <- c(210,63,111,333)
res2 <- chisq.test(tulip2, p = c(4907/6578,645/6578,451/6578,575/6578))
res2
```

```
##
##  Chi-squared test for given probabilities
##
## data:  tulip2
## X-squared = 1441.8, df = 3, p-value < 2.2e-16
```

```
# applied_race_dif
tulip <- c(22, 3,1, 19)
ind_res <- chisq.test(tulip, p = c(36/73, 1/73, 8/73,28/73))
```

```
## Warning in chisq.test(tulip, p = c(36/73, 1/73, 8/73, 28/73)): Chi-squared
## approximation may be incorrect
##  Chi-squared test for given probabilities
##
## data:  tulip
## X-squared = 12.528, df = 3, p-value = 0.005778
```

```
# applied_ai_dif
tulip <- c(22, 3,1, 19)
equal_res <- chisq.test(tulip, p = c(1/4,1/4,1/4,1/4))
equal_res
```

```
##
##  Chi-squared test for given probabilities
##
## data:  tulip
## X-squared = 31, df = 3, p-value = 8.5e-07
```

```r
# aaai_race_dif
tulip2 <- c(210,63,111,333)
eq_res2 <- chisq.test(tulip2, p = c(1/4,1/4,1/4,1/4))
eq_res2
```

```
##
##  Chi-squared test for given probabilities
##
## data:  tulip2
## X-squared = 238.53, df = 3, p-value < 2.2e-16
```