

Statistical Consulting Report

Restricted Entry: Is bias limiting Sikhs' access to mental health services?

Consultant: Ning Shen

Clients: Maria Stahre, Rajeena Kumar

Abstract

Studies have shown that in North America Racial/Ethnic (RE) disparities exist among social service providers when providing social services to minority groups [1, 2]. Researchers of this study aim to investigate RE bias within the population of counsellors in the province of British Columbia towards Punjabi Sikh individuals in terms of how receptive the service provider is to provide social services. In this statistical report, logistic regression and ordinal regression are proposed for the two response variables related to responsiveness and reception rate respectively. Potential confounding factors for one of the predictors are also discussed. Furthermore, a power analysis is included so that the researchers may decide a proper sample size for the upcoming experiment.

1 Introduction

Research has identified RE disparities in delivery of social services such as counselling in North America. Literatures review of RE biases has found that counselling service providers seem to have an implicit bias, that is, a positive attitude towards Caucasians and a negative attitude towards people of color [1, 2]. This study aims to investigate RE bias at the entry point of social services. Specifically, the study seeks to examine the possibility of RE disparities in accessing mental health services by a particular South Asian (SA) group of individuals (i.e., Punjabi Sikh individuals) in Canada. The main interest is the service providers' possible bias towards religious individuals of Sikh background in terms of the service providers' receptiveness to the service recipients. The secondary objective is to investigate the impact of other factors such as gender, accent and/or intergroup contact on the accessibility of the services.

2 Study Description

The participants of this study will be randomly selected social service providers, i.e., counsellors, from professional associations or licensure bodies common to BC as found on their public directories. Participants will receive a standardized (i.e., pre-recorded) voicemail representing one of the conditions regarding the status of religion, gender and/or accent through random assignment. A pilot study has identified that certain Sikh names and Christian names are highly likely to conjure up an image of a Sikh or Christian person, respectively. An assumption is thus made that the participants are able to identify the religion of the caller by his/her name. In terms of intergroup contact, the experiment will be performed on psychological counsellors from two cities with distinct percentage of Punjabi Sikh individuals. To reduce the possibility that the phone call will be answered, all calls will be placed after regular working hours (e.g., 8:30 pm). A pilot experiment with 20 participants will be conducted to determine the sample size required to obtain a power of 0.8 due to the scarcity of research in this area. Whether the service provider returns the phone message and offers appointments will be recorded as response variables. An example of all possible variables is shown in Table 1. In the real dataset, columns will be variables in Table 1 and rows will be phone messages.

One thing in particular to note is that, variable *city* is included in the analysis as one of the predictors instead of the level of intergroup contact because of the potential confounding factors, such as demographic characteristics of the participants in different cities. In other words, participants in one city are less prejudiced towards Sikh patients not necessarily because of a higher intergroup contact, but also possibly because of a higher proportion of Sikh counsellors in that city. Therefore, we could only make inference on the *city* factor through the analysis rather than the factor of intergroup contact. To infer the causal relationships between RE disparities and level of intergroup contact, further investigation beyond the scope of this study is needed.

Variable Type	Variable Name	Possible Values
Predictors / independent variables	<i>religion</i>	‘Sikh’ or ‘Non-Sikh’
	<i>gender</i>	‘Male’ or ‘Female’
	<i>accent</i>	‘Present’ or ‘Absent’
	<i>city</i>	‘Richmond’ or ‘Surrey’
Responses / dependent variables	<i>callback</i>	‘True’ (coded as ‘1’) or ‘False’ (coded as ‘0’)
	<i>appoint_offer</i>	‘Yes’, ‘Not sure’ or ‘No’ (coded as ‘2’, ‘1’, ‘0’ respectively)

Table 1: Example table of variables included in the analysis.

3 Proposed Statistical Analysis

Since the project is only at its early stage with no data available, simulated data are generated for a clear presentation of the proposed analysis under the assumption that the data structure and possible values in Table 1 are correct. A logistic regression model is proposed for the response coded as a binary variable (i.e., *callback*). An ordinal regression model, which is an extension of the logistic regression model, is recommended for the case where the response is coded as an ordinal variable with more than two levels. Specifically, an ordinal variable is a variable whose value exists on an arbitrary scale where only the relative ordering between different values is important, such as *appoint_offer*. A sample analysis though statistical programme language R is illustrated below as well as interpretation for the results. In the following analysis we include all predictors from Table 1, but choice of the predictors is flexible due to nature of the models. Researchers of the study may choose a desired set of predictors depending on the circumstances. All programming codes can be found in my Github repository.

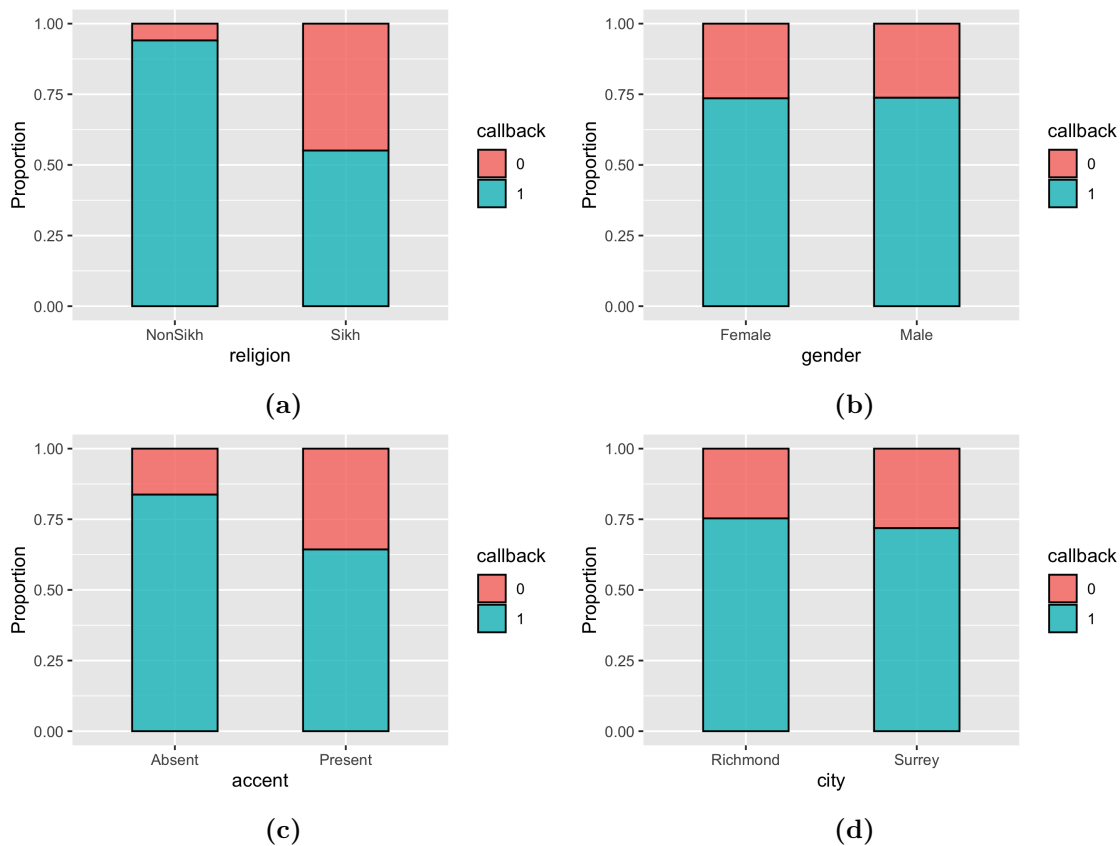


Figure 1: Distribution of *callback* in predictors a) *religion*, b) *gender*, c) *accent* and d) *city*.

3.1 Logistic regression on *callback*

As an exploratory plot, Figure 1 depicts the distribution of *callback* within possible values of the predictors respectively. *Religion* and *accent* are the two factors that have distinctive proportions of ‘no callback’ in different categories, while *gender* and *city* seem to have little influence on the callback rate. Further, we use interaction plots to determine if any possible interaction terms should be included in the model. Figure 2 is an example of the interaction plot between predictors *religion* and *accent*. The plot shows that the slopes of the lines are pretty close, i.e. the decrease of callback rate from ‘non-Sikh’ to ‘Sikh’ does not differ by the status of *accent*, thus no interaction term is needed between *religion* and *accent* in the regression model. We also provide two simulated situations in the Appendix Figure A.1, where the slopes of the lines differ drastically, indicating the two predictors have interaction effect present. Examination of interactions between every pair of predictors shall be done by the researchers before fitting the model.

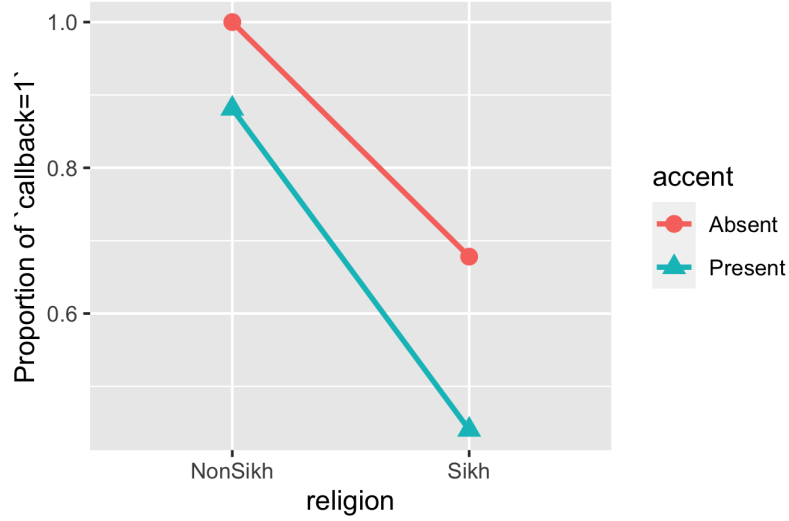


Figure 2: Interaction plot between *religion* and *accent*. Y-axis is the proportion of ‘*callback=1*’ in a group, and x-axis is the status of *religion*. Color represents the status of *accent*.

3.1.1 Fitting with no interaction term

Assuming there is no detected interaction between predictors, we build the logistic regression model on response variable *callback* taking four factors into account without interaction terms:

$$\log \left(\frac{\mu_i}{1 - \mu_i} \right) = \beta_0 + x_{i1}\beta_1 + x_{i2}\beta_2 + x_{i3}\beta_3 + x_{i4}\beta_4, \quad i = 1, 2, \dots, n, \quad (1)$$

where i represents the i^{th} phone message, $\mu_i = E(y_i|\mathbf{x}_i) = P(y_i = 1|\mathbf{x}_i)$ is the probability of i^{th} phone message receives callback, $\mathbf{x}_i = [x_{i1}, \dots, x_{i4}]$ are the values of predictors *religion*, *gender*, *accent* and *city* for the i^{th} phone message, β_0 is the intercept, and β_1, \dots, β_4 are the regression coefficients for four predictors respectively. For each of these binary predictors, one of the levels is taken as a baseline, and the other levels is included as a dummy variable in the model.

A main advantage of the logistic regression model is its attractive interpretation: $\frac{\mu_i}{1-\mu_i} = \frac{P(y_i=1)}{P(y_i \neq 1)} = \frac{P(y_i=1)}{P(y_i=0)}$ can be interpreted as the **odds** of event “ $y_i = 1$ ”, so the parameter (say) β_j may be interpreted as the change of odds in log-scale when the j^{th} predictor is changed from baseline value to alternate value. We can fit the regression coefficients in Equation 1 through R. The following presents a sample code output from applying the `glm` function on simulated data with the function parameters set to perform logistic regression:

```
> lr.fit <- glm(callback ~ religion + gender + accent + city,
+               data = dat, family=binomial("logit")) ## Fit the parameters
> summary(lr.fit) ## Results summary
```

Call:

```
glm(formula = callback ~ ., family = binomial("logit"), data = dat)
```

Deviance Residuals:

	Min	1Q	Median	3Q	Max
	-2.2947	-0.9460	0.2638	0.7694	1.4280

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	3.76833	0.35272	10.684	< 2e-16 ***
religionSikh	-2.70256	0.28465	-9.494	< 2e-16 ***
genderMale	-0.02224	0.21612	-0.103	0.9180
accentPresent	-1.21003	0.22274	-5.432	5.56e-08 ***
citySurrey	-0.40560	0.21796	-1.861	0.0628 .

From the output, the estimated intercept is 3.77 and the estimated regression slope for the four predictors are -2.70, -0.02, -1.21, -0.41 respectively. The output gives us the fitted model:

$$\log \frac{P(\text{callback} = 1 | \text{religion}, \text{gender}, \text{accent}, \text{city})}{P(\text{callback} = 0 | \text{religion}, \text{gender}, \text{accent}, \text{city})} = 3.77 - 2.70 * I(\text{religion} = \text{'Sikh'}) - 0.02 * I(\text{gender} = \text{'male'}) \\ - 1.21 * I(\text{accent} = \text{'present'}) - 0.41 * I(\text{city} = \text{'Surrey'}), \quad (2)$$

where $I(\cdot)$ is the indicator function.

Take *religion* for an example of interpretation. Here the baseline value is set as “non-Sikh” thus the alternate value is “Sikh” accordingly. The exponentiated estimate of coefficient for “religionSikh” in the R output, $e^{-2.70} = 0.067$, is the factor by which the odds of receiving callback is multiplied when *religion* changes from “non-Sikh” to “Sikh” while other variables remain fixed. In other words, the odds of a voicemail from a Sikh person receiving callback is expected to reduce $1 - 0.067 = 93.3\%$ compared with a non-Sikh person. Correspondingly, we could observe from Figure 1a that the callback rate in the “Sikh” group is notably lower than the “non-Sikh” group. The other three predictor coefficients can be interpreted similarly. Meanwhile, the exponentiated intercept, $e^{3.77} = 43.4$, represents the odds that a phone message receives callback when it is from a non-Sikh female with no accent in city Richmond, i.e., the baseline.

3.1.2 Fitting with interaction terms

If interaction effects between variables are detected (say a 2-way interaction between *religion* and *accent*), we build the model with the detected interaction terms:

$$\log\left(\frac{\mu_i}{1 - \mu_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i1} x_{i3}, \quad i = 1, 2, \dots, n, \quad (3)$$

where β_5 is the regression coefficient for the interaction term, and the rest of the notations are identical with Equation 1. The example R code is shown below:

```
> lr.fit <- glm(callback ~ religion + gender + accent + city + religion:accent,
+               data = dat, family=binomial("logit"))
```

To be noted, adding an interaction term to a model drastically changes the interpretation of the coefficients. If there were no interaction term, β_1 would be interpreted as the unique effect of *religion* on the log-odds of *callback*. But the interaction means that the effect of *religion* on *callback* varies for different values of *accent*. The unique effect of *religion* is represented by everything that is multiplied by *religion* in the model: $\beta_1 + \beta_3 * I(\text{accent}=\text{'present'})$. β_1 is now interpreted as the unique effect of *religion* on the log-odds of *callback* only when $I(\text{accent}=\text{'present'}) = 0$.

3.1.3 An alternative regression on the callback rate

Because all explanatory variables are discrete factors, an alternative regression analysis to the logistic regression may be built on the callback rates (i.e., the proportions of “*callback*=1”) in all possible combinations of the predictors. For instance, there will be a total of 16 combinations

if *religion*, *gender*, *accent* and *city* are all included. However, rather than using the proportions directly as the response, we suggest applying a logit transformation to the callback rate so that the domain of the response is expected to be the real line rather than 0 to 1. Denoting r as the callback rate, the logit transformation is given by

$$\text{logit}(r) = \log\left(\frac{r}{1-r}\right). \quad (4)$$

An example regression equation with no interaction term would be

$$\text{logit}(r_j) = \beta_0 + \beta_1 x_{j1} + \beta_2 x_{j2} + \beta_3 x_{j3} + \beta_4 x_{j4}, \quad j = 1, 2, \dots, d, \quad (5)$$

where r_j is the callback rate from the j^{th} combination, $\mathbf{x}_i = [x_{j1}, \dots, x_{j4}]$ are the values of *religion*, *gender*, *accent* and *city* from the j^{th} combination, and β_1, \dots, β_4 are the regression coefficients. Though this regression equation looks similar to Equation 1, interpretation of the coefficients is different. Here, the estimated coefficient β_j of a predictor is interpreted as the expected difference in $\text{logit}(r)$ between that level of the variable and its baseline given other predictors fixed. For example, if *religion* changes from “non-Sikh” to “Sikh”, i.e., from baseline to the alternative level, the logit-transformed callback rate will increase a value equal to β_1 . The corresponding R code can be found here.

3.2 Ordinal regression on *appoint_offer*

Instead of the odds of being one category in logistic regression, the ordinal regression model focuses on the odds of being less than or equal to a particular category. In our case, this odd is defined as $\frac{P(Y \leq k)}{P(Y > k)}$ for $k = 0, 1$ where Y is the response variable *appoint_offer*, and k is the category in the response. Here we won't go into the details of the case with interaction effects, but researchers shall resemble the procedure of adding interaction terms as in subsection 3.1. Here the ordinal logistic regression model is thus parameterized as

$$\log \frac{P(y_i \leq k | \mathbf{x}_i)}{P(y_i > k | \mathbf{x}_i)} = \beta_{k0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4}, \quad i = 1, 2, \dots, n, \quad k = 0, 1, \quad (6)$$

where notations are the same as Equation 1 except that the intercept β_{k0} varies with category k . We use the `polr` function in MASS package from R to estimate the parameters in the regression model. The sample code runs as follows:

```
> or.fit <- polr(appoint_offer ~ religion + gender + accent + city, data = dat, Hess=TRUE)
> summary(or.fit)
Call:
```

```
polr(formula = appoint_offer ~ religion + gender + accent + city,
     data = dat, Hess = TRUE)
```

Coefficients:

	Value	Std. Error	t value
religionSikh	-3.6144	0.2379	-15.1946
genderMale	-0.3360	0.1838	-1.8281
accentPresent	-2.3898	0.2131	-11.2122
citySurrey	-0.1545	0.1844	-0.8377

Intercepts:

	Value	Std. Error	t value
0 1	-5.5780	0.3543	-15.7447
1 2	-2.9841	0.2705	-11.0338

According to the output, the estimated model can be written as:

$$\log \frac{P(\text{appoint_offer} \leq 0 | \text{religion}, \text{gender}, \text{accent}, \text{city})}{P(\text{appoint_offer} > 0 | \text{religion}, \text{gender}, \text{accent}, \text{city})} = -5.58 - 3.61 * I(\text{religion} = \text{'Sikh'})$$

$$- 0.34 * I(\text{gender} = \text{'male'}) \quad (7)$$

$$- 2.39 * I(\text{accent} = \text{'present'})$$

$$- 0.15 * I(\text{city} = \text{'Surrey'})$$

and

$$\log \frac{P(\text{appoint_offer} \leq 1 | \text{religion}, \text{gender}, \text{accent}, \text{city})}{P(\text{appoint_offer} > 1 | \text{religion}, \text{gender}, \text{accent}, \text{city})} = -2.98 - 3.61 * I(\text{religion} = \text{'Sikh'})$$

$$- 0.34 * I(\text{gender} = \text{'male'}) \quad (8)$$

$$- 2.39 * I(\text{accent} = \text{'present'})$$

$$- 0.15 * I(\text{city} = \text{'Surrey'})$$

The estimated regression slope for the four predictors are the same for the two odds. Interpretations for the coefficients are similar to that of the logistic regression model (Eq. 1), except for the intercepts. The exponent of first intercept $e^{-5.58} = 0.004$ represents the odds that a baseline voicemail receives “Yes” as the answer for requesting appointment offer. The exponent of second intercept $e^{-2.98} = 0.051$ represents the odds that a baseline voicemail receives “Yes” or “Not sure” as the answer for requesting appointment offer. In this model output, we could observe that *religion* and *accent* have larger influence on the possibility of receiving appointment offer, which is in accordance with Figure 3. Furthermore, similar to Section 3.1.2, interaction terms shall be

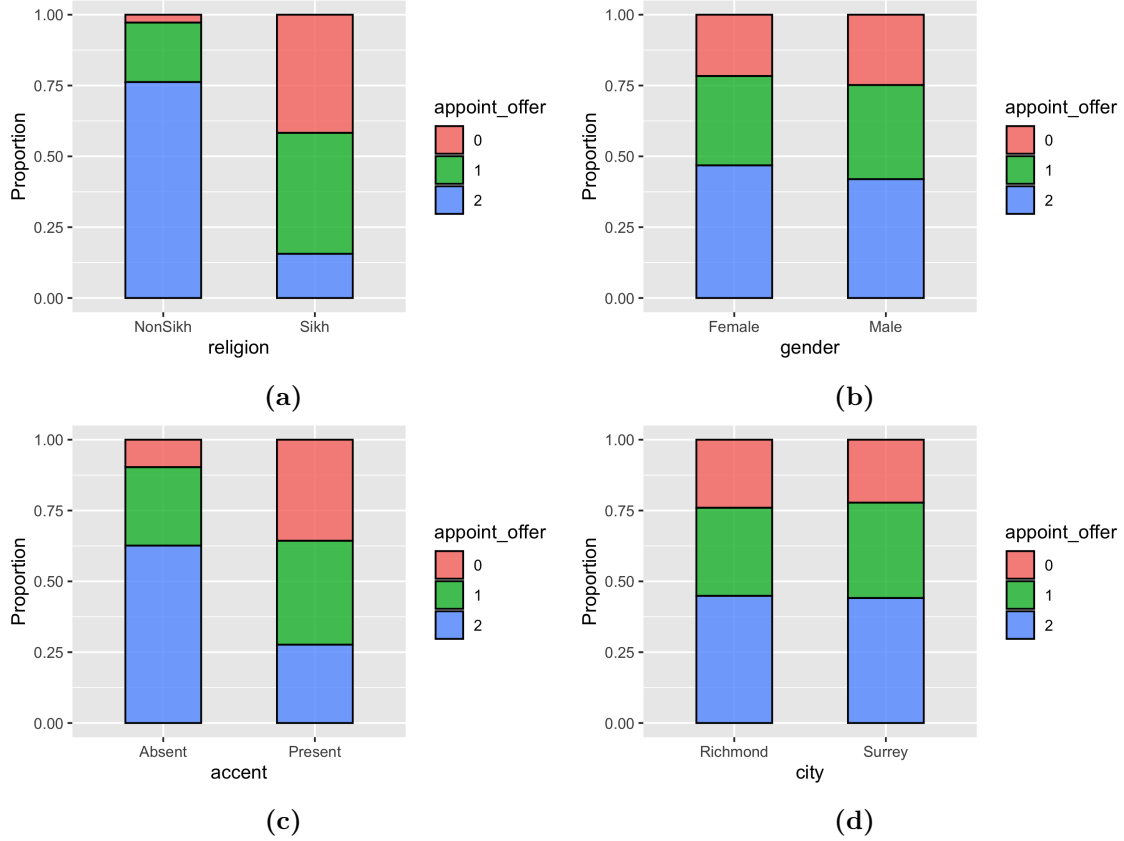


Figure 3: Distribution of *appoint_offer* in predictors a) *religion*, b) *gender*, c) *accent* and d) *city*.

added to the model if they are suggested from the exploratory plots.

4 Sample Size Calculation

As mentioned above, a pilot experiment with 20 participants will be conducted for estimation of an appropriate sample size. A power analysis, which allows us to determine the sample size required to detect an effect of a given size with a given significance level, is included in this section. The statistical power of a hypothesis test is the probability of detecting an effect, given that there is a true effect present. On the contrary, the significance level is the probability of finding an effect that is not there. Since the main objective of the study is associated with religious disparities, we focus on predictor *religion* and determine the sample size based on 2-sample comparison of proportions. There are two response variables included in the study, but for simplicity we demonstrate the analysis using *callback*. The approach for power analysis introduced by Wang and Chow [3] is adopted in this analysis. For programming details, please refer to my Github repository.

The information that need to be gathered from the pilot experiment is the proportion of 1's in *callback* given that *religion* is 'non-Sikh', $P(Y = 1|X = \text{'non-Sikh'}) \triangleq p_1$, and the proportion of 1's in *callback* given that *religion* is 'Sikh', $P(Y = 1|X = \text{'Sikh'}) \triangleq p_2$, where *callback* is denoted as Y and *religion* is denoted as X . Then the following formula is used to calculate the total sample size N for the study:

$$N = 2 \cdot \frac{p_1(1 - p_1) + p_2(1 - p_2)}{(p_1 - p_2)^2} \cdot (z_{\alpha/2} + z_{\beta})^2, \quad (9)$$

where α is the significance level, $1 - \beta$ is the statistical power, $z_{\alpha/2}$ is the critical value of the Normal distribution at $\alpha/2$ (e.g. for a confidence level of 95%, α is 0.05 and the critical value is 1.96), z_{β} is the critical value of the Normal distribution at β (e.g. for a power of 80%, β is 0.2 and the critical value is 0.84), and p_1 and p_2 are the expected sample proportions of the two groups defined above.

Figure 4 is a power analysis plot with power calculated by the abovementioned formula. It shows the statistical power with a significance level of 0.01, 0.05 and 0.1 under sample size constraints given $p_1 = 0.60$, $p_2 = 0.40$. Sample size corresponding to different values of p_1 and p_2 can be obtained through applying the formula by researchers themselves.

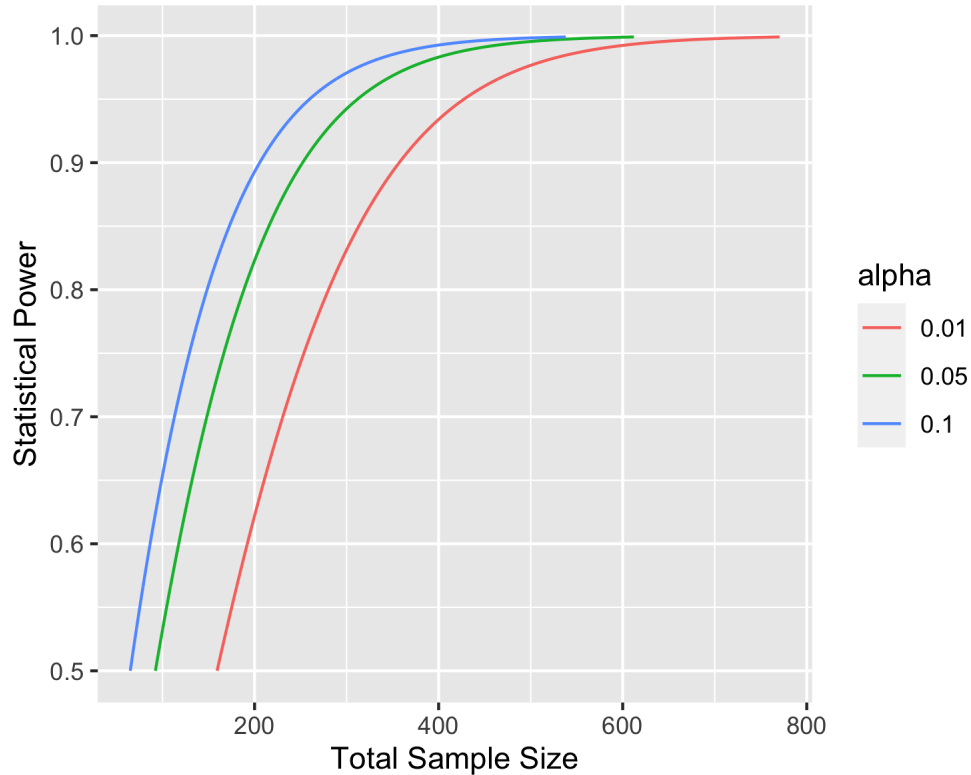


Figure 4: Power analysis plot for 2-proportion comparison with significance level of 0.01, 0.05 and 0.1.

5 Conclusion

To explore the potential bias from counsellors towards Sikh individuals at the entry point of the service, we recommend a logistic regression and an ordinal regression model to estimate the effect of factors like religion and gender on the counsellors' responsiveness and reception rate. Sample analyses for the regression models are demonstrated in the report and researchers may perform the analysis as depicted on a flexible choice of interested factors. However, the researchers should remain cautious with the factor of intergroup contact as it might be confounded with other factors such as demographic characteristics of the counsellor. An example power analysis for the sample size determination is also presented.

A Figures

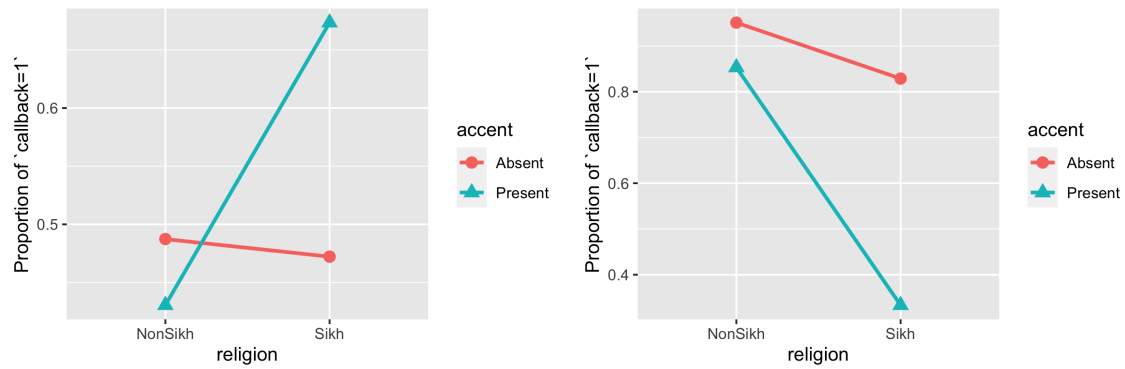


Figure A.1: Examples of interaction plot between *religion* and *accent* with interaction present. Y-axis is the proportion of 'callback=1' in a group, and x-axis is the status of *religion*. Color represents the status of *accent*.

References

- [1] Heather Kugelmass. “sorry, i’m not accepting new patients”: An audit study of access to mental health care. *Journal of Health and Social Behavior*, 57(2):168–183, 2016. doi: 10.1177/0022146516647098. URL <https://doi.org/10.1177/0022146516647098>. PMID: 27251890.
- [2] Richard Q. Shin, Lance C. Smith, Jamie C. Welch, and Ijeoma Ezeofor. Is allison more likely than lakisha to receive a callback from counseling professionals? a racism audit study. *The Counseling Psychologist*, 44(8):1187–1211, 2016. doi: 10.1177/0011000016668814. URL <https://doi.org/10.1177/0011000016668814>.
- [3] Hansheng Wang and Shein-Chung Chow. *Sample Size Calculation for Comparing Proportions*. American Cancer Society, 2007. ISBN 9780471462422. doi: <https://doi.org/10.1002/9780471462422.eoct005>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/9780471462422.eoct005>.