Report Title: Example Report

Group 3
Department / Course

November 15, 2025

Contents

1 Executive Summary

In the UK it is a criminal offence to drive a motor vehicle with a blood or breath alcohol concentration above the prescribed limit. When a person is arrested for driving under the influence of alcohol it is not usually possible to perform an accurate test of the level of alcohol in the blood or breath immediately. Breath tests can be used as an initial screening tool at the scene, but these are not sufficiently accurate for prosecution. Instead, people are taken to a police station or hospital, where the test can be carried out using proper laboratory protocols. As the body clears alcohol from the blood through time this means that if the individual was over the limit, the measured blood alcohol concentration(BAC) will be lower at the time of measurement than it was when the person was driving a motor vehicle. To deal with this situation, If the BAC after time t (hours) is measured as C_t (g/kg), the BAC at time 0 is estimated as $C_0 = C_t - \beta t$, where β $(g/kg/h)$ is BAC elimination rate.

The key point is how to find a precise β to estimate C_0. Forensic scientists currently 2.5% percentile of β distribution constructing from samples as the estimated β value for every individuals. This method is obviously not rigorous enough for the courts since:

- The courts will be forced to make decisions under estimated β if we only give a single estimation of *beta*, since the calculated C_0 is either over or under the legal limit.
- Differences between individuals are ignored, for example age and sex, which may affect β.
- β value at 2.5% is over conservative.

2 Motivations: Need improve? Why?

3 Disadvantage of current method

1. beta is fixed for all individuals, we need to consider the heterogeneity in individuals.

2. Deviated aim: Police Scotland aims at assessing the probabilities that a person is over the limit, while the parameter beta only concentrates on the range between 2.5% and 97.5%.

4 New method: build a linear regression model for beta

Check if age, weight and height have linear relationships with beta. Or have any patterns?

5 Model selection:

use drop(), look into AIC and R^2

6 Hypothesis test(70yr female example-better use 2.5th quantile?):

H0: C0 < limit, H1: C0 > limit

7 Calculate V_d:

in the Widmark's equation assumption, beta and V_d are independent, so we need to check whether Cov(V_d, beta) = 0

8 Further research:

casual effect: biased data selection, (e.g. inner correlations between sex and height)

8.1 REPORT

9 Exclusive summary

background Goal - 1, 2, 3 main data + elimination rate explanation method - model we used result

10 Overview of Datasets

10.1 Variable Description

- how many observations do we have? where we get the resourse, reference? explain each variables in table ## Prelimary analysis
- initial formula: explain beta60 + - motivation to improve?

variable correlations? Why? hypothesis needed? why 2.5 quantiles?

11 Model selection regarding key variable 'beta60'

11.1 Advantage of current method

simpler model with only one parameter(beta), easier to calculate, suitable for no dataset situations

11.2 New model motivation

Current model disadvantage: 1. beta is fixed for all individuals, we need to consider the heterogeneity in individuals. 2. Deviated aim: Police Scotland aims at assessing the probabilities that a person is over the limit, while the parameter beta only concentrates on the range between 2.5% and 97.5%.

12 Bayesian model

intro New model (Bayesian) ## Model Selection and Evaluation - why Bayesian? Any advantage, any improvement? More accurate? … ##explanation ###why prior setted? ## how result reflect?

12.1 Model fitting example

How our new model works on the 70 year-old madam?

13 Testing another assumption(condition):

13.1 variable explanation

explain all variables we used explain needed references

##Testing whether formula reasonable? whether beta60 relate to Vt? how related? hows the result? how new model fit?

#Limitation our model limit data collected reason - cannot reflect the real world model reason other reason: variables