



## Saturday Morning Statistics #24

May 14<sup>th</sup>, 2022



# Enter the Skunk



*“‘It’s been hard to breathe and the smell’s been just horrendous. [It’s] like you’ve literally been sprayed by a skunk.’”*

- Resident of Prague, Oklahoma

*“‘It’s nasty’: Prague neighbors push back on area cannabis facility, Oklahoma News 4*

<https://kfor.com/news/local/its-nasty-prague-neighbors-push-back-on-area-cannabis-facility/>

## Over eight hundred cannabis strains characterized by the relationship between their psychoactive effects, perceptual profiles, and chemical compositions

by Laura Alethia de la Fuente, Federico Zamberlan, andres sanchez, Facundo Carrillo, Enzo Tagliazucchi, Carla Pallavicini

- **Claim:** *“While cannabinoid content was variable even within individual strains, terpene profiles matched the perceptual characterizations made by the users and could be used to predict associations between different psychoactive effects.”*

<https://www.biorxiv.org/content/10.1101/759696v1.abstract>

# Question and Hypothesis

## Question of the day.

- Can we build a model to predict a cannabis product's **potential** effects and flavors given the product's lab results.



## Methodology: Probit Models

Given a latent variable representation of the **probit model**:

$$z_i = x_i\beta + \epsilon_i, \quad \epsilon_i \stackrel{iid}{\sim} \mathcal{N}(0, 1),$$

$$y_i = \begin{cases} 1 & \text{if } z_i > 0 \\ 0 & \text{if } z_i \leq 0 \end{cases}$$

You can estimate the parameters using the **likelihood function**

$$L(\beta) = \prod_{i=1}^n \Phi(x_i\beta)^{y_i} [1 - \Phi(x_i\beta)]^{1-y_i}.$$



**Thank you for coming.**

### Insights of the Day

- Stand on the shoulders of giants to reach the hard-to-get fruit.
- Simple is better than complex.