Prosem 2020: Project 1: Signal Detection

Your Name Here, University of Colorado Boulder

?? September 2020

Part I: E.P.T Pregnancy Test

EPT Data

Table 1: Number of 'no' and 'yes' EPT responses for pregnant (s1) and not-pregnant (s0) women in the 1977 clinical trials of the Warner-Lambert Early Pregnancy Test.

EPT Results	Not Pregnant (s0)	Pregnant (s1)
no	183	36
yes	15	451

Transform EPT Data

EPT ROC Graphs and Model Graph

Part 2: Axelsson (2018) Experiment 1

Dual-Gaussian, Equal-Variance, Single-Criterion Model ("not sick or"sick")

Axelsson and colleagues (Axelsson et al., 2018) asked subjects to judge whether or not photographs of a person show a person infected with influenza virus or not infected with the virus. In Study 1 of their paper 62 subjects viewed individual photographs of 16 (8 men; 8 women) people who, two hours earlier, had been injected either with an influenza virus (*Escherichia coli endotoxin*), the **LPS** condition or with saline (0.9% NaCl), the **placebo** condition. The subjects were asked to decide if each photograph was of a "sick" person or a "healthy person". Two composite photos from their study, each made up of the average of eight women two hours after injection of placebo (right) and the same 8 women after injection of the virus (left) are shown in Figure 3.

Read Study 1 Data

Table 2 gives the number of "healthy" and the number of "sick" judgments made by 62 subjects shown the 32 individual photographs one at a time. Each subject saw both sick and placebo photos of the 16 people (for a total of 32 photographs) in a random order with the constraint that no photograph of the same person was shown immediately following the other photo of that person.

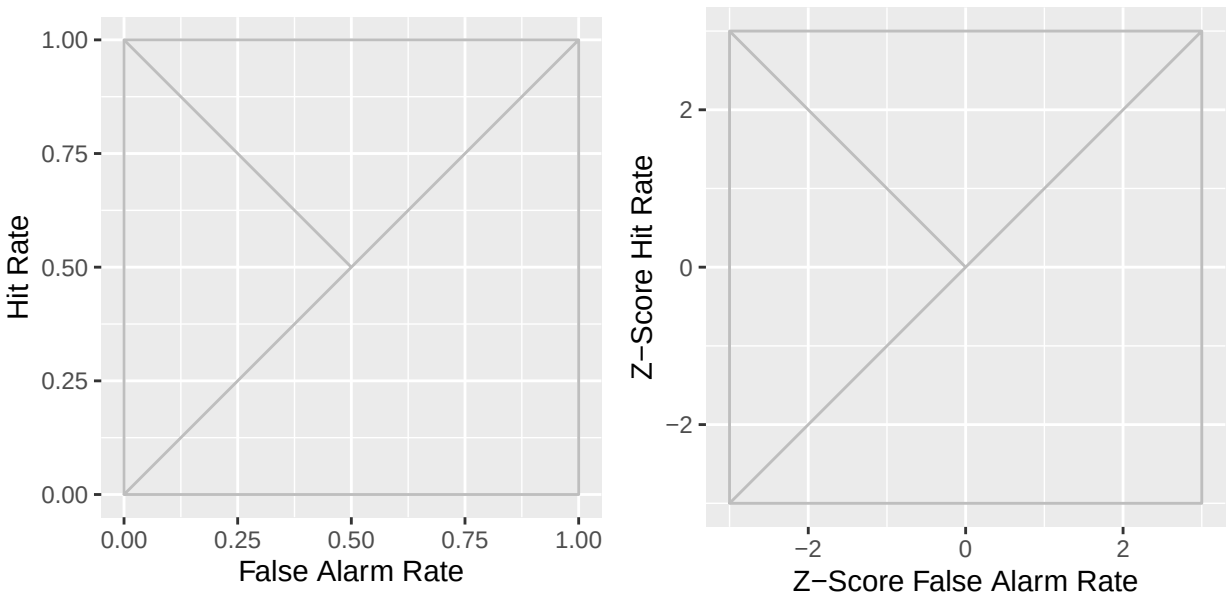

Figure 1: Receiver operating characteristic of the Warner-Lambert E.P.T. home pregnancy test. The filled circle is the resulting hit rate and false alarm rate computed from the clinical trials data. The smooth curve represents the predictions of the equal-variance signal detection model. The gray positive diagonal represents the hit rates and false alarm rates that would occur if the test had no ability to predict pregnancy. The gray negative diagonal represents the hit rates and false alarm rates that could occur with an unbiased test.

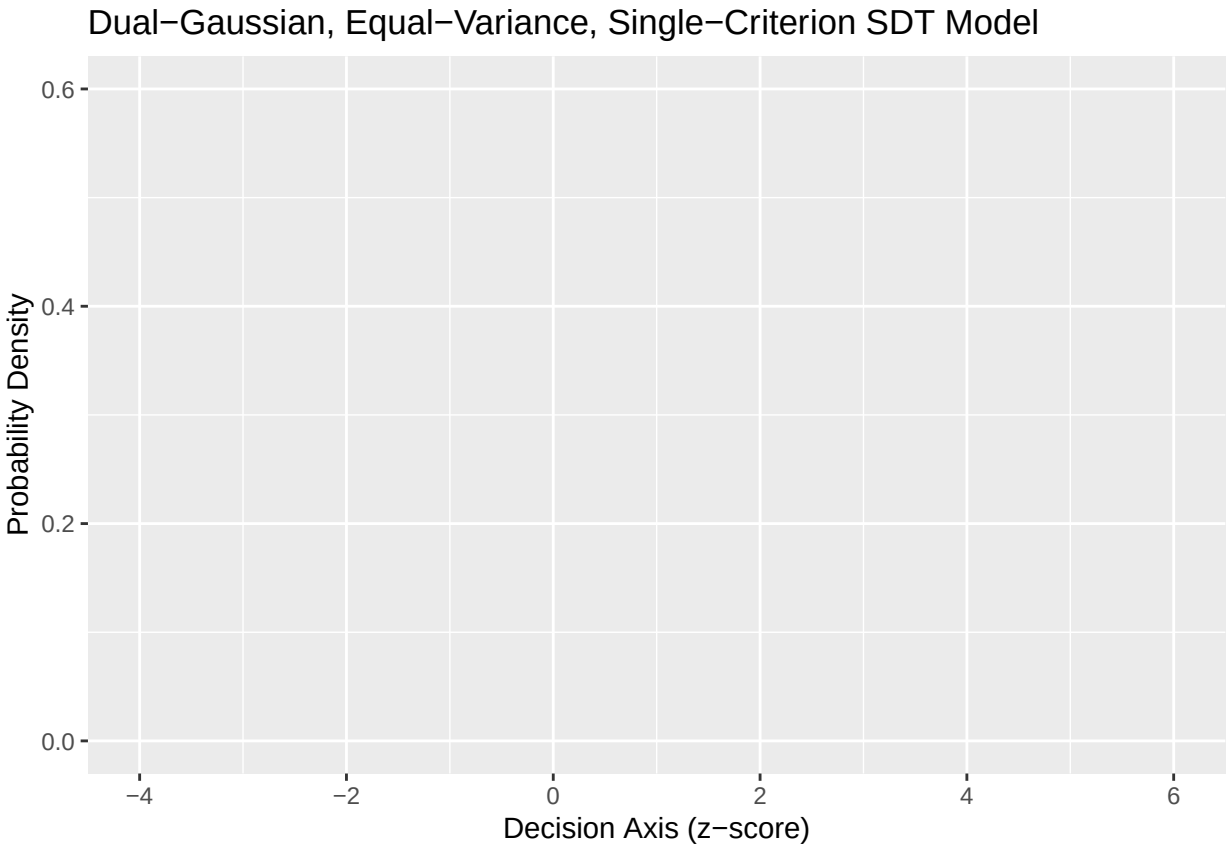

Figure 2: The dual-Gaussian, equal-variance, single decision criterion model of the Warner-Lambert E.P.T. pregnancy test. The red distribution represents the test output distibution in response to non-pregnancy; the blue distribution represents the test output distribution to pregnancy. The vertical black line is the decision criterion. It divides the response space into 'non-pregnant' and 'pregnant' decision outcomes.

Figure 3: Figure 3 from Axelsson et al., 2018. Which composite photo, a or b, is of an infected person?

Table 2: The number of 'healthy' and the number of 'sick' judgments made by 62 subjects shown the 32 individual photographs (16 infected with the lps virus and 16 injected with a placebo saline solution) one at a time. The data are from Axelsson et al. (2018). The frequencies are aggregated across all subjects.

Rating	Placebo (s0)	Virus (s1)
healthy	1028	702
sick	440	775

Make Graphs

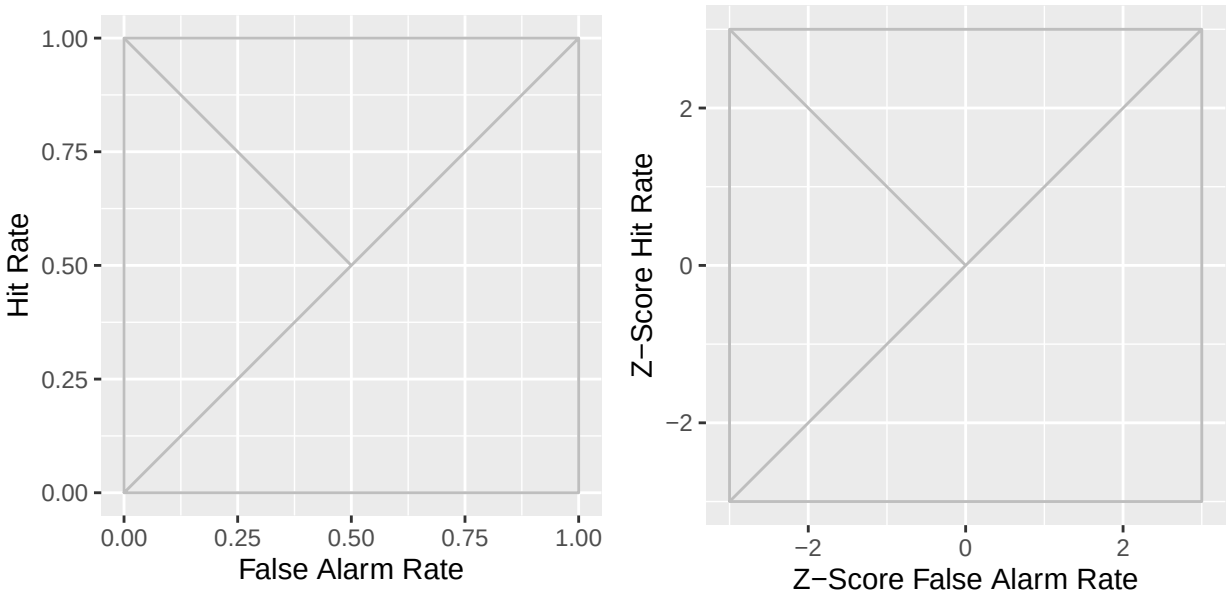

Figure 4: Receiver operating characteristic of subjects judging photographs as 'sick' or 'healthy' computed from the data of Study 1 in Axelsson, et al (2018). The left panel plots the probability; the right panel plots the z-score of the probability. The filled data point is the resulting hit rate and false alarm rate for detecting the sick photos, computed from the aggregated data of all subjects. The smooth curve represents the predictions of the equal-variance signal detection model. The gray positive diagonal represents the hit rates and false alarm rates that would occur if there was no ability to detect the sick photos. The gray negative diagonal represents the hit rates and false alarm rates that could occur with an unbiased observer.

Part 3: Unequal-Variance, Multiple Decision Criteria Model (Sick People)

Axelsson et al. (2018) reported a second study using these photographs. They asked 60 new subjects to rate the 32 photographs on how sick the person appeared to be using a confidence rating scale that ranged from R1 ("very poor") to R7 ("very good"). Table 3 shows the number of times each rating was given for each type of photograph. Since we are considering this task to be one of detecting the sick person, I have reversed the order of the ratings, making R1 mean "very healthy") and R7 mean "very sick" as presented in Table 3.

Table 3 gives the number of times each confidence rating was used for each image condition:

Table 3: The number of times for each condition that the subjects rated the photograph using a 1 to 7 confidence rating scale. The frequencies are aggregated across all subjects.

Rating	Placebo (s0)	Virus (s1)
R1	617	476
R2	2917	1799
R3	3451	3032
R4	3025	2892
R5	2109	2784
R6	673	1640
R7	81	294

Transform Data

Graphs of the Model

Conclusions

What do you conclude about peoples' ability to judge health from photographs? How does the d' and accuracy (A_z) from Study 1 compare with d_a and A_z from Study 2? How do these subjects compare to the Warner-Lambert E.P.T. test?

References

Axelsson, J., Sundelin, T., Olsson, M. J., Sorjonen, K., Axelsson, C., Lasselin, J., & Lekander, M. (2018). Identification of acutely sick people and facial cues of sickness. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1870).