

GR5224 Homework 1 Answer Key

1 Gambling

Answers certainly could vary, but here are a few important points.

The use of probability in blackjack is close to the Frequentist perspective, where the probability of an event is defined as the proportion of times that it would happen in the limit as the number of hands played approaches infinity. It is valid to calculate a *conditional* probability, which is the essence of various schemes that count the number of cards that have previously been dealt.

One relatively simple approach discussed in the video is to count small cards as +1, medium cards as 0, and high cards as -1. When the sum of this count divided by the number of remaining decks is sufficiently negative, the expected value of playing becomes (slightly) positive — primarily due to the increased chance of getting a blackjack which pays 1.5 times the bet — so it makes sense to increase the bet size compared to when the count is near zero. When the count divided by the number of remaining decks is positive, it makes sense to reduce your bet size to the table minimum. You expect to lose money but need to burn through enough low cards to get to a state where the count is more favorable.

If the count and the number of remaining decks is known, the Fisher would not have an objection to using conditional probability this way. However, once you introduce the possibility that the player may make a mistake when counting, then the count could be considered a random variable whose expectation is what the player believes it is and the spread depends on the rate that the player makes mistakes. At that point, Fisher would object to the use of probability to describe what a person believes, since this is inconsistent with the above definition of probability, differs from one person to the next, and makes probability subjective rather than a sole property of the deck being dealt from.

In poker, the deck is shuffled after every hand, so there is no counting but you can condition on the cards that are visible to you. Poker is the rare example where objective and subjective probability are simultaneously active. You can objectively calculate the probability of various events, given the cards that are visible to you. Although your opponent's cards are not visible to you, you can subjectively condition on their betting behavior as well as past hands that you have played against them. Knowing the objective probabilities is utterly insufficient to play poker well because what you really need to know the probability that your hand ends up being better than your opponent's.

The person who invented game theory (around the same time as Fisher) used poker as a unifying [metaphor](#) but embraced the subjective view of probability:

Von Neumann was only interested in poker because he saw it as a path toward developing a mathematics of life itself. He wanted a general theory — he called it “game theory” — that could be applied to diplomacy, war, love, evolution or business strategy. But he thought that there could be no better starting point than poker: “Real life consists of bluffing, of little tactics of deception, of asking yourself what is the other man going to think I mean to do. And that is what games are about in my theory.”

Thus, it is no surprise that Bayesian updating is central to game theory, but it is surprising that Bayesian estimation is so rare outside of statistics departments.

One of the most interesting parts of the video starts about 21:40 in where Galen Hall asserts “An introduction of a modest amount of gambling into most people’s lives, for many people would actually be a net positive, because it would teach them to think about how to weight uncertain outcomes ...” To which, Liv Boeree responds “It gives people statistical literacy ... anything that trains people to understand probabilities and just think through things with that, like with a greater degree of nuance and granularity ... is a service to humanity.” Decision theory with posterior probabilities is a belief management system. Although the Bayesian approach is subjective in the sense that one person could make a different decision than another — even if they observe the same data — due to differences in prior probability distributions and / or utility functions, it is very systematic rather than arbitrary and can be applied to any situation where you need to make a decision without knowing everything that you would like to know with certainty. That certainly includes investing other people’s money, but also normal business operations, personal finance, and science.

2 Economic Growth

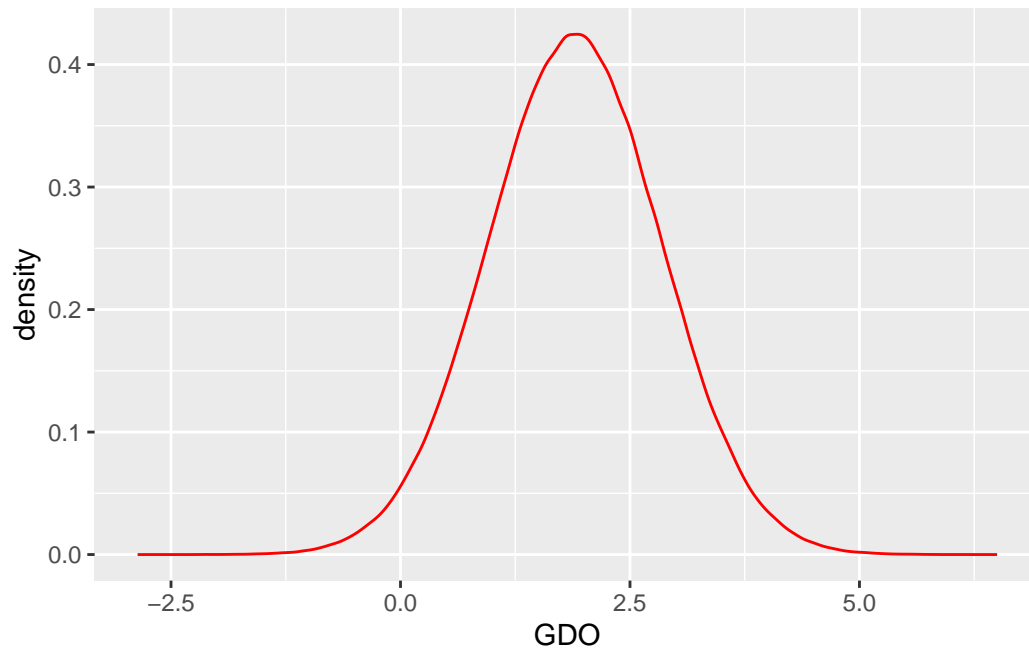
2.1 Prior Distribution

```
library(dplyr)
m <- 1.9
s <- 0.8
R <- 10^6
sims <- tibble(mu = rnorm(R, mean = m, sd = s))
```

The value of m is the average of the economist’s forecasts for the first quarter of 2025 (as of mid-November 2024) and the value of s is their typical root mean-squared error when forecasting the next quarter.

2.2 Prior Predictive Distribution

```
sigma <- 0.7
library(ggplot2)
sims <- mutate(sims,
               GDP = rnorm(R, mean = mu, sd = sigma),
               GDI = rnorm(R, mean = mu, sd = sigma),
               GDO = (GDP + GDI) / 2)
ggplot(sims) + geom_density(aes(x = GDO), color = "red")
```



GDO has more uncertainty than does the prior on μ , due to the measurement noise which has a standard deviation of 0.7. However, GDO has less uncertainty than do either of GDP or GDI due to the averaging.

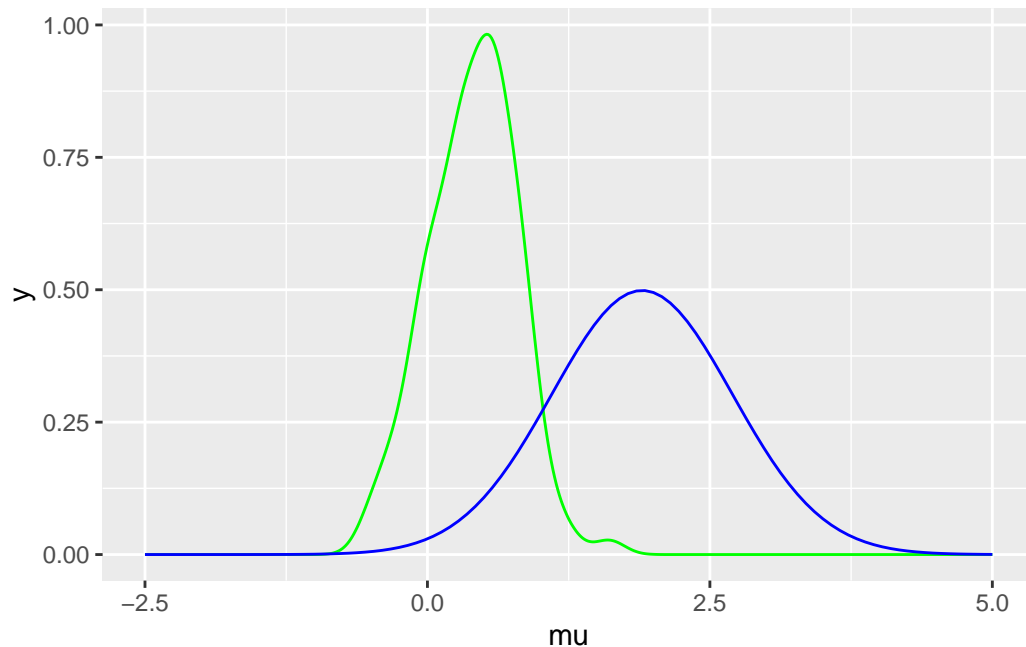
2.3 First Quarter of 2025

```
post <- filter(sims,
  round(GDP, digits = 1) == -0.5,
  round(GDI, digits = 1) == 0.2)
nrow(post) / R
```

```
[1] 0.000111
```

The proportion of simulations we keep after conditioning on the observed data is an estimate of the denominator of Bayes Rule, which in this case is the probability of observing — under your model for the data-generating process — GDP growth being -0.5 and GDI growth being 0.2 . As is often the case, this number is rather small even with few data points.

```
ggplot(post) +
  geom_density(aes(x = mu), color = "green") +
  xlim(-2.5, 5) +
  geom_function(fun = dnorm, args = list(mean = m, sd = s), color = "blue")
```



The posterior density of μ in green is shifted to the left and concentrated relative to the prior density in blue, due to conditioning on the observed data. However, the posterior distribution does not imply that you are sure that μ is zero. Rather, you are pretty sure it is between -1 and 1 , with a center that is just slightly above zero.

2.4 Second Quarter of 2025

Answers could vary, but the Wall Street Journal released a survey of economists today

```
ROOT <- "https://prod-i.a.dj.com/public/resources/documents/"
FILE <- "wsjecon0725.xlsx"
download.file(paste0(ROOT, FILE), destfile = FILE)
WSJ <- readxl::read_excel(FILE, skip = 1)
mutate(WSJ, mu = as.numeric(`Second Quarter 2025`)) |>
summarize(median = median(mu, na.rm = TRUE),
           IQR = IQR(mu, na.rm = TRUE))
```

```
# A tibble: 1 x 2
  median    IQR
  <dbl> <dbl>
1   2.13   1.02
```

Note that the Inter-quartile range (IQR) is the difference between the economist at the 75th percentile and the economist at the 25th percentile. The uncertainty of these two economists (or any other ones) is not reflected in the WSJ survey, but is presumably similar to the value of for s , namely 0.8 , used above.