# STAT 165/265 HW 9

March 28, 2025

**Due Friday, April 4 at 11:59pm**

*Expected completion time: 120 minutes*
*Graded on accuracy*

## Practice: Squiggle

**Supplemental reading:** Squiggle Documentation

In lecture 12, we saw how to use the software tool *Squiggle* to automatically incorporate numerical sensitivity and structural uncertainty into our forecasts. Last homework, we simulated the number of days until the Omicron hospital peak using Python. In this exercise, we will practice doing this in Squiggle, then apply the same techniques to a new forecasting problem. We will begin with a review of the basic syntax for Squiggle.

For each of these problems, feel free to test your code in a fresh Squiggle playground environment.
**IMPORTANT:** Write up your solutions as you go, as your code may not be saved.

1) Syntax Basics
   **Supplemental reading:** Creating Distributions in Squiggle

   Review this demo of how to make basic distributions and do numerical computations in Squiggle. Then, answer the following questions:

   a) What distribution does this Squiggle code snippet simulate?
      **Hint:** Consider its shape. The answer is also in the Squiggle documentation (linked above)

      ```
      q1_a = 100 to 900
      ```

   b) Write a piece of Squiggle code to simulate a $\mathcal{N}(\mu = 95, \sigma^2 = 25)$ distribution. Include a screenshot of your simulated distribution in your submission.

c) Write a piece of Squiggle code to simulate a LogNormal distribution with 5th and 95th percentiles as 0.05 and 0.25 respectively. Include a screenshot of your simulated distribution in your submission.

d) Write a piece of Squiggle code to simulate a distribution whose *logarithm* is Normal with mean 0.05 and standard deviation 0.25. Include a screenshot of your simulated distribution in your submission.

e) Write a piece of Squiggle code to create a mixture of a $\mathcal{N}(\mu = 50, \sigma^2 = 100)$ and a point mass at 95 with relative weights of 19 and 1, respectively.

f) Suppose we want to make a LogNormal distribution based on a 95% confidence interval. Write a Squiggle function called dist95 that returns a LogNormal distribution, given 2.5th and 97.5th percentiles. Test your function to generate a LogNormal distribution for the 95%confidence interval [165, 2650]. Include both your code *and* a screenshot of your simulated distribution in your submission.
**Hint:** If instead we wanted to make a LogNormal distribution based on a $(1 - \alpha)\%$ confidence interval, we would need to change "factor" in the dist70 function to be the value

$$\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$$

so that we have $1 - \alpha/2$ area to the left of "factor" on a standard normal distribution. Here, $\Phi^{-1}(\cdot)$ denotes the inverse standard normal CDF. You may use an online calculator to find the new factor.

2) Modifying the Omicron Hospitalization Peak Forecast

In lecture 12, we saw a demonstration of how to simulate our forecast for the number of days until the peak for Omicron hospitalizations using Squiggle. Recall the formula we came up with and the definition for its parameters:

$$T = \log_2 (N/2N_0) \cdot t + \Delta_0 + \Delta_1$$

where:

- $N_0$ is the total number of future UK Omicron cases
- $N$ is the current number of UK Omicron cases
- $t$ is the Omicron doubling time
- $\Delta_0$ is the lag between case peak and hospital peak
- $\Delta_1$ is the lag between single-day hospital peak and 7-day average hospital peak

and our beliefs for each were:

- $N_0 \in [150K, 250K]$ with 70% confidence.
- $N \in [5M, 13M]$ with 70% confidence.
- $t \in [2.0, 3.3]$ with 70% confidence.
- $\Delta_0 + \Delta_1 \in [9, 14]$ with 70% confidence.

We also had the following structural uncertainties:

- Herd immunity: 15% chance of +5 days (28 million vs. 7 million cases)

- Short doubling time: 15% chance of 1.5 day doubling instead of 2.4
- Extended peak: 10% chance of +[0,12] days

Suppose some news comes out that changes our beliefs. Based on the news, we want to modify the forecast we made earlier. We now believe:

- $N_0 \in [150K, 250K]$ with 70% confidence (no change)
- $N \in [8M, 15M]$ with 70% confidence (new lower and upper bounds)
- $t \in [2.2, 3.14]$ with 95% confidence (new lower and upper bounds, higher confidence)
- $\Delta_0 + \Delta_1 \in [8, 15]$ with 95% confidence (new lower and upper bounds, higher confidence)
- Extended peak: 20% chance of +[0,10] days (higher chance, max of 10 extra days instead of 12)

Suppose we also wish to change the confidence level on our forecast to be **80%**, rather than 70%. **Hint:** to do this, we don't need to make any additional changes the confidence levels on the individual components. We can simply change which quantiles we take from the final distribution. For example, if our final distribution was stored as dist_final and we wanted a 90% confidence interval we would take:

```
lower = quantile(dist_final, 0.05)
upper = quantile(dist_final, 0.95)
```

so that we have 5% mass to the left of the lower bound and 5% mass to the right of the upper bound.

Implement these updated beliefs in Squiggle and change the confidence level to 80% to generate a new forecast for $T$, the number of days until the peak for UK Omicron hospitalizations. Express your final answer in terms of the number of days from December 21, 2021.

Include your code, final answer (in number of days), and a screenshot of the simulated distribution for the number of days until the peak in your submission.

3) Applying Squiggle to New Forecasts

a) Numerical Sensitivity

Suppose we have decomposed a forecasting problem for a quantity of interest, $\gamma$ into the following parts:

$$\gamma = \alpha * \beta + \Delta$$

where we have some uncertainty about each of $\alpha, \beta, \Delta$:

- $\alpha \sim \mathcal{N}(\mu = 85, \sigma^2 = 16)$
- $\beta \sim \mathcal{N}(\mu = 10, \sigma^2 = 25)$
- $\Delta \sim \text{Unif}([0, 1])$

Use Squiggle to find an 80% confidence interval for $\gamma$.

b) Structural Uncertainty

Suppose we have decomposed a forecasting problem for a quantity of interest into two "worlds."

- World 1 (88% chance): $\gamma \sim \mathcal{N}(\mu = 25, \sigma^2 = 4)$
- World 2 (12% chance): $\gamma = 15$

Use Squiggle to find an 80% confidence interval for $\gamma$.

## Predictions

*Expected completion time: 90 minutes*
*Graded on accuracy as part of the class forecasting competition*

Make and submit predictions to the questions on this Google Form:
https://forms.gle/KMdFqtxakV2xuTdV9.

Be sure to follow the format described at the top of the form. For each question, you will submit a mean and inclusive 80% confidence interval or a probability (whichever the question asks for). We provide cells on the Google form for you to type out your reasoning (1-2 paragraphs), which you should submit to Gradescope with the rest of this assignment. For questions 1-3, your prediction (but not the explanation) will appear on the public leaderboard.

## [STAT 265 only] None this week