

RETURN PERIODS IN JULIA

An aerial photograph of a residential neighborhood in Julia, Mississippi, completely inundated with floodwater. The water is a murky brown color, reaching the roofs of many houses and surrounding the trees. The houses have various roof colors, including red, grey, and brown. The trees are mostly green, and some are partially submerged. A central road runs vertically through the middle of the image, also flooded. In the upper right corner, a green golf course is visible, partially surrounded by water. The overall scene depicts a severe flooding event.

James Doss-Gollin

CEVE 543

2026-08-26

Waiting and risk

Return periods from data

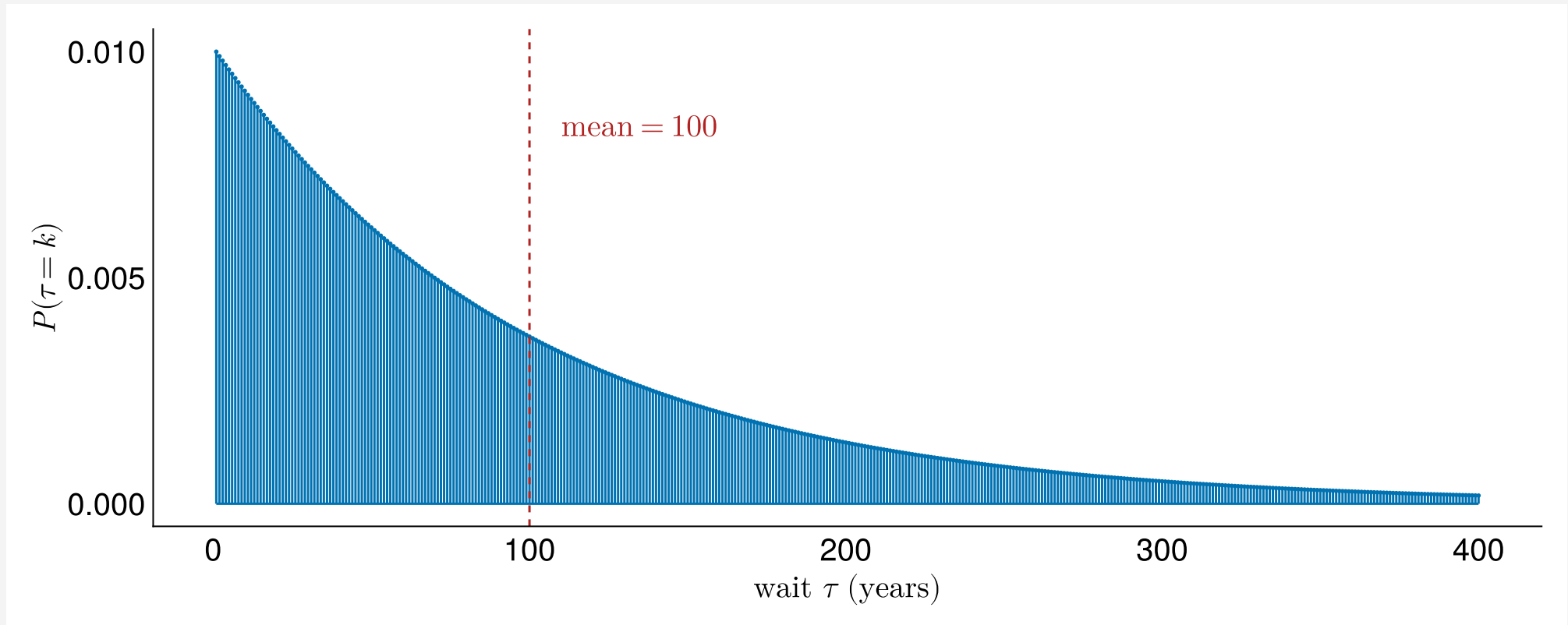
WAITING TIME FOR THE NEXT EXCEEDANCE, BY HAND

τ is the number of years until the “coin” next comes up “heads”. By the geometric distribution:

$$p_{\tau}(k) = (1 - p)^{k-1} p, \quad E[\tau] = \frac{1}{p} = T$$

WAITING TIME FOR THE NEXT EXCEEDANCE, BY SIMULATION

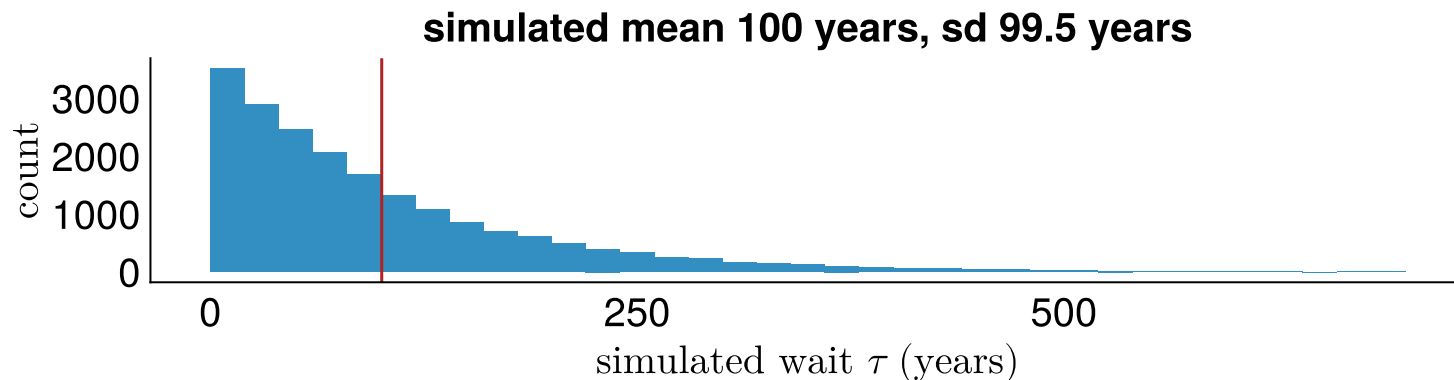
$$p = 0.01$$



HOW VARIABLE IS THAT WAIT?

$$\text{Var}(\tau) = \frac{1-p}{p^2}, \quad \text{sd}(\tau) = \frac{\sqrt{1-p}}{p} \approx T \text{ for small } p$$

The spread is as large as the mean wait, so the mass near 1 and the long tail are the same picture.



HOW LIKELY ARE TWO EXCEEDANCES IN TEN YEARS, AT $p = 1/4$?

Ten flips of the bent coin, so $N = 10$, with Y the number of heads.

$$p_Y(y) = \binom{N}{y} p^y (1 - p)^{N-y}$$

```
1 Y = Binomial(10, 0.25)
2 p_two_or_more = 1 - pdf(Y, 0) - pdf(Y, 1)
3 @printf("P(Y >= 2) = %.3g", p_two_or_more)
```

```
P(Y >= 2) = 0.756
```

A “4-year flood” turns up at least twice per decade more often than not.

RARER EVENTS: $p = 1/10$

Same ten flips, same count Y , a coin bent one tenth of the way instead of one quarter.

```
1 Y = Binomial(10, 0.1)
2 p_two_or_more = 1 - pdf(Y, 0) - pdf(Y, 1)
3 @printf("P(Y >= 2) = %.3g", p_two_or_more)
```

```
P(Y >= 2) = 0.264
```

We get two “ten-year floods” in a decade a quarter of the time.

WHAT RISK DOES A DESIGN LIFE CARRY?

$$R = 1 - (1 - p)^N, \quad T = \frac{1}{1 - (1 - R)^{1/N}}$$

```
1 risk(T, N) = 1 - (1 - 1 / T)^N
2 T_needed(R, N) = 1 / (1 - (1 - R)^(1 / N))
3
4 @printf("100-year flood, 50-year life: R = %.3g\n", risk(100, 50))
5 @printf("10%% risk over a 5-year job needs T = %.3g years", T_needed(0.10, 5))
```

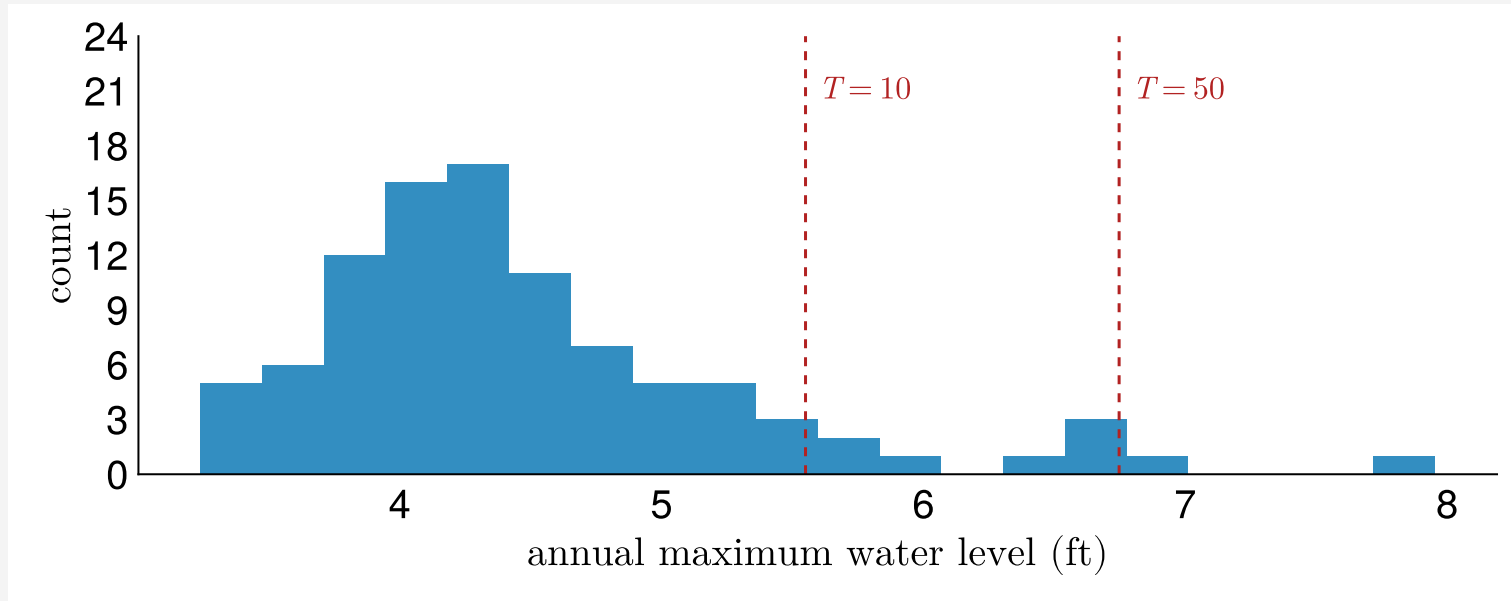
```
100-year flood, 50-year life: R = 0.395
10% risk over a 5-year job needs T = 48 years
```

Waiting and risk

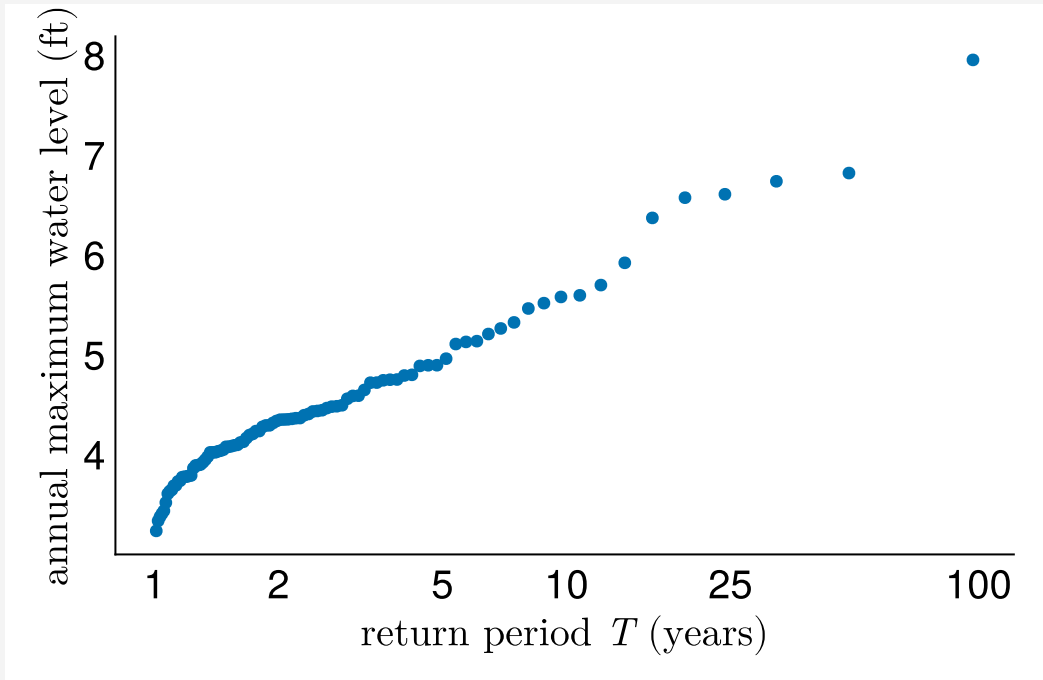
Return periods from data

A HISTOGRAM HIDES THE TAIL

Histograms make it hard to compare and visualize tails.



WHAT DOES A RETURN PERIOD PLOT SHOW?



Which axis is data, and which is an estimate?

- Level: a number a gauge recorded.
- Return period: computed from the rank of that level in a finite record.

HOW IS THE RETURN PERIOD ESTIMATED?

Sort descending, so rank $i = 1$ is the largest, and a hat marks a quantity computed from the record rather than known.

$$\hat{p}_i = \frac{i}{n+1}, \quad \hat{T}_i = \frac{1}{\hat{p}_i}$$

Over n years an event with annual probability p is exceeded about np times, so a value exceeded i times out of n estimates $\hat{p} \approx i/n$. Dividing by $n+1$ instead gives the Weibull plotting position.