# Decision and Risk Analysis

## Lecture 8: **Posterior Inference & Prediction**

DAEN 427 / ISEN 427 / ISEN 627 - Fall 2026

**Alexander Abuabara**



# Learning objectives

- Summarize a posterior with a point estimate and credible interval.
- Evaluate hypotheses using posterior probabilities and Bayes factors.
- Explain and compute posterior predictions for new obs.
- Use MCMC draws to approximate and check posterior summaries.

# Our example: artists at MoMA

- Let $\pi$ be the fraction of major modern-art museum artists born in 1965 or later.
- In the MoMA sample, $Y = 14$ of $n = 100$ artists meet this definition.
- Prior: $\pi \sim \mathrm{Beta}(4, 6)$.
- Sampling model: $Y \mid \pi \sim \mathrm{Bin}(100, \pi)$.

$$\underbrace{\mathrm{Beta}(4,6)}_{\text{prior}} \xrightarrow{Y=14 \text{ of } 100} \underbrace{\mathrm{Beta}(18,92)}_{\text{posterior}}.$$

The data move our estimate toward 14%,
while the prior still contributes information.

# R: inspect the sample and update

```
library(bayesrules)
library(tidyverse)
data("moma_sample")

moma_sample %>%
  group_by(genx) %>%
  tally()

plot_beta_binomial(alpha = 4, beta = 6,
                   y = 14, n = 100)
```

- `bayesrules`: provides `moma_sample` and plotting helpers.
- `tidyverse`: `group_by()`, `tally()`, and the %>% pipe.
- `genx`: indicator for an artist born in 1965 or later.
- `plot_beta_binomial()`: passes prior shapes $4, 6$ and observed $y = 14$, $n = 100$.

The posterior is concentrated near 0.16
and much narrower than the prior.

# Why is a point estimate not enough?

**Question:** What percentage of artists are Gen X or younger?

$$E(\pi \mid Y = 14) = \frac{18}{110} \approx 0.164,$$
$$\text{Mode}(\pi \mid Y = 14) = \frac{17}{108} \approx 0.157.$$

- A typical posterior value is about 16%.
- But a point estimate hides uncertainty:
  - $\text{Beta}(18, 92)$ vs. a wider $\text{Beta}(4, 16)$ posterior based on 10 observations. (check the book!)
- Two analyses can have similar centers but very different precision.

# A 95% posterior credible interval (CI)

Let $f(\pi \mid y)$ be the posterior density, and $\pi_q$ its $q$th quantile.

$$\text{Middle 95\% CI} = (\pi_{0.025}, \pi_{0.975}),$$

$$P(\pi_{0.025} < \pi < \pi_{0.975} \mid Y = 14) = 0.95.$$

For the posterior, the interval is approximately $(0.10, 0.24)$.

## In plain-language ...

Given the model and observed data, there is 95% posterior probability that the population proportion is in this interval.

# R

```r
# Middle 95% credible interval
qbeta(c(0.025, 0.975), 18, 92)
# [1] 0.1009 0.2379

# Middle 50% credible interval
qbeta(c(0.25, 0.75), 18, 92)
# [1] 0.1388 0.1862

# Middle 99% credible interval
qbeta(c(0.005, 0.995), 18, 92)
# [1] 0.0853 0.2647
```

- `qbeta()` is a *base R* function: it returns Beta distribution quantiles.
- First argument: quantile probabilities for the lower and upper endpoints.
- Arguments `18, 92`: posterior Beta shape parameters.
- The 50% interval is narrowest; the 99% interval is widest.

## Intuition

Higher credibility requires including more posterior values, usually widening the interval. Highest-density intervals are an alternative to equal-tailed intervals.

# Posterior hypothesis testing

## One-sided question: is $\pi < 0.20$?

$$H_0 : \pi \geq 0.20 \qquad H_a : \pi < 0.20.$$

The probability of the claim is an area under the posterior density:

$$P(H_a \mid Y = 14) = \int_0^{0.20} f(\pi \mid Y = 14)\, d\pi = 0.849.$$

Given the prior, sampling model, and 100 observed artists, the posterior chance that the population proportion is below 20% is about 84.9%.

# R (probability, odds, and Bayes factor)

```r
post_prob <- pbeta(0.20, 18, 92)
post_odds <- post_prob / (1 - post_prob)

prior_prob <- pbeta(0.20, 4, 6)
prior_odds <- prior_prob / (1 - prior_prob)
BF <- post_odds / prior_odds

post_prob  # 0.849
post_odds  # 5.622
prior_prob # 0.08564
prior_odds # 0.09366
BF         # 60.02
```

- `pbeta(x,a,b)` (base R): $P(\pi \leq x)$ for a $\mathrm{Beta}(a, b)$ variable.
- `post_odds`: probability of $H_a$ divided by probability of $H_0$ after observing data.
- `prior_odds`: the same ratio before observing data.
- `BF`: posterior odds divided by prior odds.

$$\mathrm{BF} = \frac{P(H_a \mid y)/P(H_0 \mid y)}{P(H_a)/P(H_0)} \approx 60.$$

## Intuition

The sample multiplies the odds for $\pi < 0.20$ by about 60;
a Bayes factor is not itself a probability.

# What about a two-sided claim?

**Question:** Is the true proportion different from 0.30?

$$H_0 : \pi = 0.30 \qquad H_a : \pi \neq 0.30.$$

- Under the continuous Beta prior and posterior, $P(\pi = 0.30) = 0$.
- So the previous odds-ratio calculation cannot test this exact point null.
- The 95% CI $(0.10, 0.24)$ lies below 0.30.
- Chapter 8 suggests a specified meaningful range: $H_0 : \pi \in (0.25, 0.35)$.

## Note

Define what counts as a meaningful difference *before* examining the data. (a point-mass prior would be a different model)

# Posterior prediction

## Predict 20 new artworks

Let $Y'$ count younger artists in 20 *new* artworks.

$$Y' \mid \pi \sim \mathrm{Bin}(20, \pi), \quad P(Y' = y' \mid \pi) = \binom{20}{y'} \pi^{y'} (1-\pi)^{20-y'}.$$

- Even for a known $\pi$, the new count varies by random sampling.
- We are also uncertain about $\pi$ itself.

### Intuition

Prediction must account for both sampling variability *and* posterior uncertainty about the rate.

# Posterior predictive distribution

Average new-data probabilities over plausible parameter values:

$$f(y' \mid y) = \int_0^1 f(y' \mid \pi) f(\pi \mid y)\, d\pi.$$

Here $y = 14$ and $\pi \mid y \sim \mathrm{Beta}(18, 92)$.

- Exact predictive probability: $P(Y' = 3 \mid Y = 14) = 0.2217$.
- $P(Y' \geq 5 \mid Y = 14) = 0.233$.
- $E(Y' \mid Y = 14) = 3.273$; the most likely count is 3.

## Meaning

Three is the best single-count prediction, not a guarantee.

# Calculate predictive probabilities

```
# Beta-binomial predictive probabilities;
# base R's beta and choose functions
ys <- 0:20
pred <- choose(20, ys) *
  beta(18 + ys, 92 + 20 - ys) /
  beta(18, 92)

pred[ys == 3]            # about 0.2217
sum(pred[ys >= 5])       # about 0.233
sum(ys * pred)           # about 3.273
```

This is an R implementation;

the book displays the analytic calculation rather than this code.

# Exact prediction

- `ys <- 0:20`: enumerate all possible counts in the next 20 observations.
- `choose(20, ys)`: count ways each outcome can occur.
- `beta()` (base R): evaluates the Beta function; its ratio averages over the posterior for $\pi$.
- `pred`: probability of each count; it sums to 1.
- `sum()` and indexing: add probabilities for $Y' \geq 5$, or calculate a weighted average.

## Intuition

The predictive distribution spreads probability across many possible future counts, not only 3.

# When exact calculations are unavailable

**Question:** Can a sample of plausible $\pi$ values replace difficult integrals?

- MCMC generates dependent draws that approximate the posterior distribution.
- Chapter 8 uses four chains with 10,000 iterations each and discards the first 5,000 of each.
- That leaves $4 \times 5{,}000 = 20{,}000$ posterior draws.
- Check simulation quality *before* interpreting summaries.

## Intuition

A posterior probability can be approximated by the fraction of simulated draws satisfying a claim.

# Specify and simulate the model

```
library(rstan)
art_model <- "
data { int<lower=0,upper=100> Y; }
parameters { real<lower=0,upper=1> pi; }
model {
  Y ~ binomial(100, pi);
  pi ~ beta(4, 6);
}"

art_sim <- stan(model_code = art_model,
                data = list(Y = 14),
                chains = 4, iter = 10000,
                seed = 84735)
```

# Explaining Stan

- `rstan` interfaces R with Stan's posterior simulation engine.
- `data`: $Y$ is the observed count, constrained between 0 and 100.
- `parameters`: unknown proportion $\pi$ is constrained to $[0, 1]$.
- `model`: binomial sampling model plus Beta prior.
- `stan()`: simulates four chains, each of length 10,000; default warmup discards half.

## Expected result

A fitted `stanfit` object containing 20,000 retained draws across four chains. Actual draw values depend on the software environment.

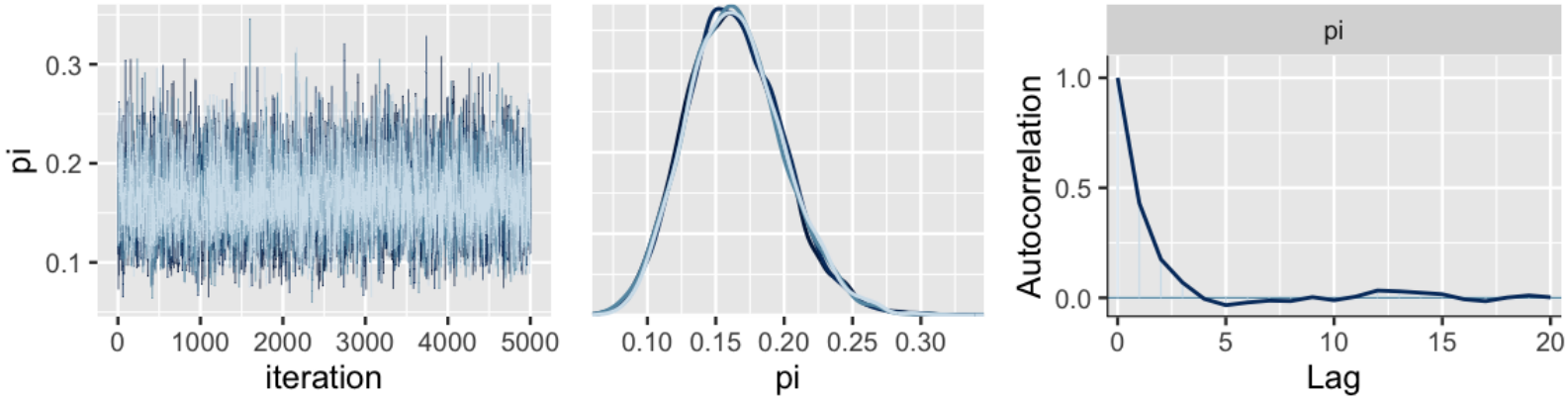
# MCMC Diagnostics

```
library(bayesplot)

mcmc_trace(art_sim, pars = "pi")
mcmc_dens_overlay(art_sim, pars = "pi")
mcmc_acf(art_sim, pars = "pi")

rhat(art_sim, pars = "pi")
neff_ratio(art_sim, pars = "pi")
```

Reported diagnostics: `Rhat = 1`; `neff_ratio = 0.378`.
Expected plots: well-mixed traces, overlapping chain densities, and quickly decaying autocorrelation (see Fig. 8.8).




- **Trace plot:** The chains fluctuate around a stable value, with no obvious trends or prolonged periods of getting stuck.
- **Posterior density:** The chains produce similar posterior distributions, centered around $\pi \approx 0.16$.
- **Autocorrelation:** The dependence between successive samples decreases quickly and approaches zero after approximately 5 lags.

The rapid decrease in autocorrelation suggests **relatively good mixing**: the chain explores the posterior efficiently rather than remaining stuck in the same region for long periods.

# Diagnostics

- `bayesplot`: trace, density overlay, and autocorrelation graphics.
- `rhat()`: checks consistency across chains; near 1 is desirable.
- `neff_ratio()`: effective sample size divided by retained sample size.
- The ratio 0.378 corresponds to approximately 7,600 independent-equivalent draws.

## Intuition

20,000 correlated MCMC draws are not necessarily as informative as 20,000 independent draws. Recheck diagnostics for your own run.

# Summarize the MCMC posterior

```
library(broom.mixed)
library(janitor)

tidy(art_sim, conf.int = TRUE,
     conf.level = 0.95)

art_chains_df <- as.data.frame(
  art_sim, pars = "pi")

mean(art_chains_df$pi)
quantile(art_chains_df$pi, c(0.025, 0.975))
mean(art_chains_df$pi < 0.20)
```

Output from `tidy()`: median 0.162, CI $(0.101, 0.239)$.

MCMC estimate of $P(\pi < 0.20 \mid y)$: 0.846.

# Simulated inference

- `broom.mixed::tidy()`: posterior median, uncertainty summaries, and central 95% interval for the fitted parameter.
- `as.data.frame()`: combine retained chain draws of $\pi$ into an R table.
- `mean()` and `quantile()`: approximate posterior center and interval endpoints.
- `mean(pi < 0.20)`: fraction of draws consistent with the claim.
- `janitor` is used for the "optional" `tabyl()` count table; it is not required for these shortened lines.

MCMC approximates the exact 95% CI $(0.1009, 0.2379)$ and exact posterior probability 0.849.

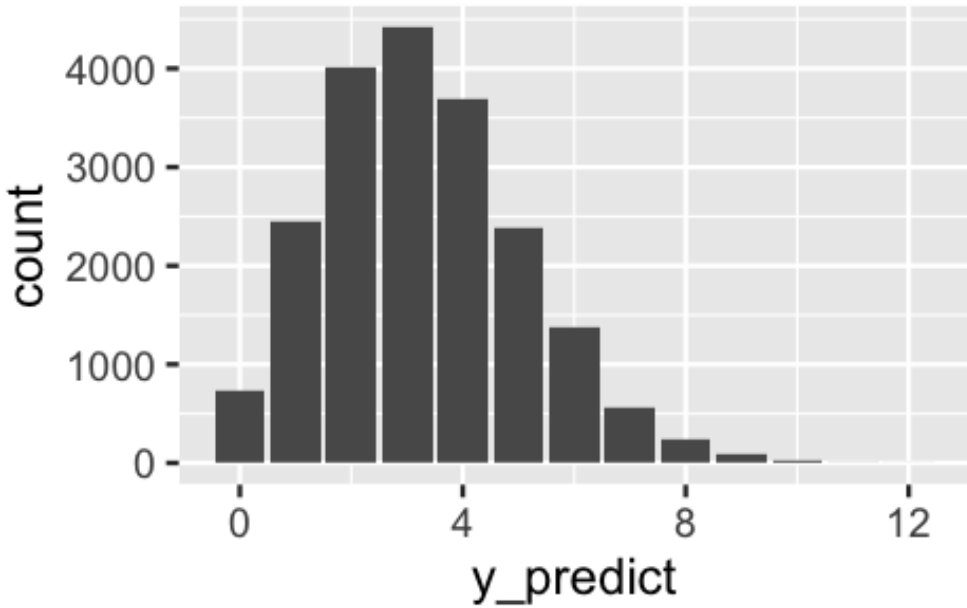
# Simulate and summarize predictions

```
set.seed(1)
art_chains_df <- art_chains_df %>%
  mutate(y_predict = rbinom(
    length(pi), size = 20, prob = pi))

ggplot(art_chains_df,
       aes(x = y_predict)) + stat_count()

art_chains_df %>%
  summarize(mean = mean(y_predict),
    lower_80 = quantile(y_predict, 0.10),
    upper_80 = quantile(y_predict, 0.90))
```

Output: mean 3.283, central 80% interval from 1 to 6.




- A histogram of 20,000 simulated posterior predictions of the number among the next 20 artists that will be Gen X or younger.
- We expect roughly 3 of the next 20 artists to be Gen X or younger, but there's an 80% chance that this figure is somewhere between 1 and 6 artists.

# Explaining the code: predictive simulation

- `set.seed(1)`: makes the R random-number stream reproducible.
- `rbinom()`: for *each* simulated $\pi$, draws a new count out of 20.
- `mutate()`: adds these counts as `y_predict`.
- `ggplot() + stat_count()`: plots the frequency of each predicted count.
- `summarize()` and `quantile()`: predictive mean and middle 80% endpoints.

## Meaning ...

Given the model and data, the simulation places about 80% of predicted 20-artwork counts between 1 and 6.

# Bayesian benefits

## Conditioning matters: posterior vs. $p$-value

Two probabilities in the chapter answer *different* questions:

$$\underbrace{P(\pi < 0.20 \mid Y = 14) = 0.849}_{\text{uncertainty about a parameter given data}}$$

$$\underbrace{P(Y \leq 14 \mid \pi = 0.20) = 0.08}_{\text{data variability under an assumed rate}}$$

### Intuition

The second quantity does not say there is an 8% chance that the assumed parameter value is true. Neither quantity is a universal true/false decision rule.

# Key takeaways

1. **Estimation:** report both a posterior center and uncertainty interval.
2. **Testing:** posterior probability assesses a claim; a Bayes factor compares changes in odds.
3. **Prediction:** average over uncertain parameters and future sampling randomness.
4. **MCMC:** approximate all three tasks from posterior draws, after checking diagnostics.

# Check your understanding

1. **Is a 95% credible interval a 95% range for a new count?**

2. **Does a Bayes factor of 60 mean $P(H_a \mid y) = 0.60$?**

3. **Why not always predict exactly 3 new younger artists?**

4. **Do many MCMC draws guarantee correct results?**

# Check your understanding (short answers here)

1. **Is a 95% credible interval a 95% range for a new count?**
   No. It describes $\pi$, not the future count $Y'$.
2. **Does a Bayes factor of 60 mean $P(H_a \mid y) = 0.60$?**
   No. It is the ratio of posterior odds to prior odds.
3. **Why not always predict exactly 3 new younger artists?**
   Both the future sample and $\pi$ vary.
4. **Do many MCMC draws guarantee correct results?**
   No. Convergence and mixing must be assessed.