Posterior prediction

ONE NEW READING

$$V_{\text{new}} = h^T V h + \sigma^2$$

COMPLETE EXAMPLE

Linear regression with known noise

```
%%%
A new reading from the beginner regression.
%%%
flatppl_compat = "0.1"
x = [0.0, 1.0, 2.0, 3.0]
y = [0.8, 2.1, 2.9, 4.2]
sigma = 0.5

X = colstack([ones(4), x])
Xt = transpose(X)
P0 = diagmat([1.0 / 25, 1.0 / 4])

V = inv(P0 + (Xt * X) / sigma^2)
m = V * (Xt * y) / sigma^2
coefficients = MvNormal(m, V)
x_new = 1.5
h = [1.0, x_new]

center = sum(h .* m)
v_mean = sum(h .* (V * h))
v_new = v_mean + sigma^2

beta, state = rand(coefficients, rnginit(42))
line = sum(h .* beta)
reading, next_state = rand(Normal(line, sigma),
    state)

z = probit(0.95)
width_mean = z * sqrt(v_mean)
width_new = z * sqrt(v_new)
```

Intervals at a fixed predictor

At $x = 1.5$, both distributions have mean 2.498.

Quantity	5% quantile	95% quantile
Fitted mean	2.087	2.909
New reading	1.579	3.417

Predict a whole curve

```
line_grid = beta[1] .+ beta[2] .* x
replica ~ Normal.(line_grid, sigma)
```

Use the same intercept and slope draw at every new input. Independent fresh noise at each input produces a simulated data set from that line.

ANNOTATED LINES

A Fix the design and prior 9–11

```
X = colstack([ones(4), x])
Xt = transpose(X)
P0 = diagmat([1.0 / 25, 1.0 / 4])
```

Each row of `X` is `[1, x]`. The intercept and slope have independent zero-mean Normal priors with standard deviations 5 and 2. The noise scale is known.

B Form the coefficient posterior 13–15

```
V = inv(P0 + (Xt * X) / sigma^2)
m = V * (Xt * y) / sigma^2
coefficients = MvNormal(m, V)
```

Add the data precision $X^T \frac{X}{\sigma^2}$ to the prior precision `P0`. Invert the sum to obtain the posterior covariance `V`. Its product with the data term gives the posterior mean `m`.

C Add the variance of a new reading 19–21

```
center = sum(h .* m)
v_mean = sum(h .* (V * h))
v_new = v_mean + sigma^2
```

At `x_new` the design row is `h`. The fitted mean and the new reading both have expectation `center`. Their variances differ by the independent measurement variance `sigma^2`. The fitted mean has variance `v_mean`.

D Simulate a new reading 23–26

```
beta, state = rand(coefficients, rnginit(42))
line = sum(h .* beta)
reading, next_state = rand(Normal(line, sigma),
    state)
```

Use one coefficient draw for the entire line. At the selected input, pass the next RNG state to the draw of a reading. This adds independent observation noise.

Take a pointwise interval

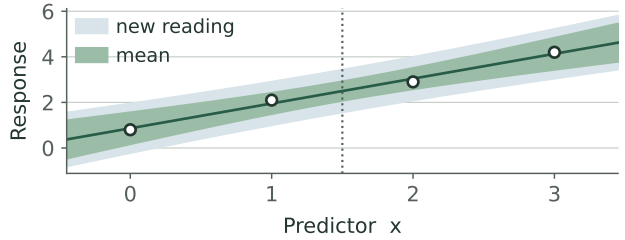
`probit(0.95)` gives the two-sided 90% Normal multiplier. Subtract and add each half-width to `center` for its corresponding interval.

Compare the readings at 0 and 3

Subtract the fitted mean at 0 from the fitted mean at 3 to get $3\beta_2$. Its posterior mean is 3.267 and its variance is $9V_{22} = 0.4425$. For two new readings at these inputs, add $2\sigma^2 = 0.5$ to the variance.

POSTERIOR AND PREDICTIVE INTERVALS

Mean line and new readings



The dark band contains 90% posterior probability for the mean at each x. The light band covers 90% of the predictive law for a new reading.

Where the bands widen

The four design values run from 0 to 3. Uncertainty in the slope matters more away from the middle of that range. Both bands widen toward the edges.

Intercept–slope covariance

The fitted mean has variance $V_{11} + 2xV_{12} + x^2V_{22}$. The middle term accounts for posterior covariance between the intercept and slope.

Two points on one fitted line

For design rows h_a and h_b, shared coefficients give covariance $h_a^T V h_b$. Independent measurement errors add σ^2 to each observation variance and leave this cross-covariance unchanged.

An unknown noise scale

If `sigma` has a prior, draw the intercept, slope and noise scale jointly from the posterior. Use one such draw for each replicated data set.

Compare observations with replications

Replicate the four observations at their original predictor values. Compare the observed maximum absolute residual, for example, with the same statistic in the replications.

Jessica Cox · bcox@mpp.mpg.de