**Step 0 – Define the goal**

We will compare

1. Prospect theory baseline: choices predicted purely from the PT values difference between risky (50/50) and sure options.
2. PT + time/repetition terms: allow Trial (time flow) and Block (repetition) to affect either the intercept (bias) or the value sensitivity (vis interactions)

**Step 1 – Load & Check Data**

# why standardisation

* Fairness for sparse selection (we should better keep the scales the same)
* Numerical stability, especially with interaction/quadratics
* Better interpretability
* Lower collinearity

# why do we measure gambling behaviour by logistic function

* Binary outcomes need probabilities: our data are 0/1 choices. To fit the model by maximum likelihood (Bernoulli), we need a valid probability *P* (gamble) for each trial

Bernoulli maximum likelihood – the principled way to estimate a probabilistic model of binary choices, it aligns with the log-loss criterion we care about, and it is interpretable

**Step 2 – fit a prospect theory baseline**

What measurements do we use to measure the models

# Best CV log loss (cross-validated log loss / negative log-likelihood)

* Measures how good your predicted probabilities are on unseen data
* Lower is better
* Strongly penalises overconfidence mistakes

# AUC (Area Under the ROC Curve)

* Measures the ranking/discrimination: how well the model ranks gambles you took (1) above gambles you skipped (0), across all possible thresholds
* Higher is better (0.5 = random; 1.0 = perfect)

# why not R square

* R square assumes squared-error to a continuous target with roughly normal, constant-variance noise. In logistic choice models the target is Bernoulli with variance *p(1-p)* that depends on the mean, and the link is nonlinear (logit). R square’s assumptions are broken
* R square ignores probability quality. You predict probabilities (e.g., 0.7 to gambles), not point values to match (0/1). R square won’t tell you if 0.7 really happens. Log loss does, it is a proper scoring rule that rewards calibrated probabilities and heavily penalizes overconfidence mistakes

**Results interpretations**

Full Feature Library

文本

描述已自动生成

dV\_PT: fitted prospect theory params

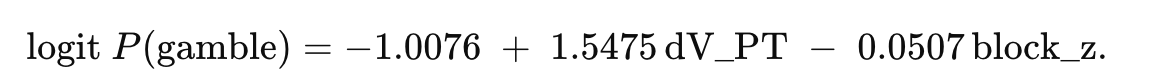
Trial\_z: z score of the variable “Trial”

Block\_z: z score of the variable “Block”

文本

描述已自动生成

So the final learnt equation is:



* From CV log loss perspective:

Baseline best CV log loss = 0.5026

Full library best CV log loss = 0.5029

Difference = + 0.0003 (full minus baseline). That’s worse (albeit by a hair), so there’s no improvement in calibrated accuracy.

* From CV AUC perspective:

Baseline best CV AUC = 0.8220

Full library best CV AUC = 0.8226

Difference = +0.0006. an tiny bump