Analyzing Response Time Data (Part 1)

Statistical Modeling in Psychology

October 9-10 2024

The Ohio State University

Outline

1. Probability Distributions
   - Rules of Probability
   - Discrete Random Variables
   - Continuous Random Variables
   - RT Distributions
2. Data from Wagenmakers et al. (2004)
3. Function Estimation
   - Cumulative Distribution Function
   - Density Function
   - Hazard Function
   - Semiparametric Estimates

Modeling

- Choosing a model
- Estimating parameters
- Evaluating fit
- Validation

Outline

The Ohio State University

Axioms

Definition (Probability)

The *probability* of an event is a measurement assigned to it that reflects the degree of plausibility or likelihood of that event. Events that are more likely to occur should have higher probabilities than events that are not as likely to occur.

Axioms

Definition (Axioms)

An *axiom* is a fundamental principle on which our measurements are based. There are three axioms of probability.

1. $0 \leq P(A) \leq 1$ for any event A. (Probability measurements are positive and represented as proportions between 0 and 1.
2. $P(\Omega) = 1$. (Something from the sample space must happen if an experiment is performed.)

Axioms

Definition (Axioms)

An *axiom* is a fundamental principle on which our measurements are based. There are three axioms of probability.

1. $0 \leq P(A) \leq 1$ for any event A. (Probability measurements are positive and represented as proportions between 0 and 1.
2. $P(\Omega) = 1$. (Something from the sample space must happen if an experiment is performed.)
3. If A and B are mutually exclusive events, then

$$P(A \cup B) = P(A) + P(B).$$

(Probability is additive over disjoint events.)

Random Variables

Definition (Random Variable)

A *random variable* is a number Y that is assigned to an event. The event may be numeric (e.g., the number of spots on a die), or it may not (e.g., the die is red). Y is a numeric code that tells us what the event was. The sample space Ω_Y lists all possible values that Y can take on.

Discrete Random Variables

Definition

Discrete Random Variable

A *discrete random variable* Y has a sample space Ω_Y that is finite or countably infinite (its elements can be mapped onto the natural numbers).

Examples

- Let Y be the number of times a coin comes up heads in 10 flips of that coin. Then $\Omega_Y = \{0, 1, 2, \ldots, 9, 10\}$. Because Ω_Y is finite, Y is discrete.
- Let Y be the number of times a coin is flipped before it comes up heads. Then $\Omega_Y = \{1, 2, 3, \ldots\}$ is countably infinite, so Y is discrete.
- Let Y be the proportion of correct responses on an exam with 100 questions. Then $\Omega_Y = \{.00, .01, .02, .03, \ldots, .98, .99, 1.00\}$ is finite, so Y is discrete.
- Note that Y does not have to take on only integer values to be discrete. However, the possible values for Y must be separated by gaps.

Discrete Univariate Distributions

Definition ($\sim$)

We will use the "$\sim$" notation to indicate a variable's distribution. For example, we might write $Y \sim \mathcal{B}(N, q)$ to say that Y follows a binomial distribution with parameters N and q.

Y	$p(y)/\mathcal{I}\ (y \in \Omega_Y)$	Ω_Y	$E(Y)$	$Var(Y)$
Geometric	$q(1-q)^{y-1}$	$\{0, 1, 2, \ldots\}$	$(1-q)/q$	$(1-q)/q^2$
Bernoulli	$q^y(1-q)^{1-y}$	$\{0, 1\}$	q	$q(1-q)$
Binomial	$\binom{N}{y} q^y (1-q)^{N-y}$	$\{0, 1, 2, \ldots, N\}$	Nq	$Nq(1-q)$
Uniform	$1/N$	$\{1, 2, \ldots, N\}$	$\bar{y}$	s_y^2
Poisson	$\mu^y e^{-\mu}/y!$	$\{0, 1, 2, \ldots\}$	μ	μ
Hypergeometric	$\dfrac{\binom{R}{y}\binom{N-R}{n-y}}{\binom{N}{n}}$	$\{0, 1, 2, \ldots, n\}$	$n\dfrac{R}{N}$	$n\dfrac{R}{N}\dfrac{(N-R)}{N}\dfrac{(N-n)}{N-1}$

Random Variables

Definition (Continuous Random Variable)

A *continuous random variable* Y has a sample space Ω_Y that is uncountably infinite (its elements consist of intervals of the real line).

Probability Density Function

Definition (Probability Density Function)

A *probability density function* (pdf) is a function that gives the *relative* likelihood (not probability) that a variable takes on a value within its sample space (*support*).

Properties

A function $f(y \mid \theta)$ is a pdf if

$$
\begin{aligned}
&1. \quad f(y \mid \theta) \geq 0 \text{ for all } y \in \Omega_Y,\\
&2. \quad \int_{y \in \Omega_Y} f(u \mid \theta) du = 1,\\
&3. \quad P(Y \in \mathcal{B}) = \int_{y \in \mathcal{B}} f(y \mid \theta) dy.
\end{aligned}
$$

Probability is represented as *area under the density function*.

Continuous Distributions

Y	$p(y)/\mathcal{I}\,(y \in \Omega_Y)$	Ω_Y	$E(Y)$	$Var(Y)$
Uniform	$1/(b-a)$	$[a,b]$	$(a+b)/2$	$(b-a)^2/12$
Exponential	$\lambda\exp(-\lambda y)$	$[0,\infty)$	$1/\lambda$	$1/\lambda^2$
Beta	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}y^{a-1}(1-y)^{b-1}$	$[0,1]$	$a/(a+b)$	$ab/[(a+b)^2(a+b+1)]$
Gamma	$\Gamma(k)^{-1}\lambda^k y^{k-1}\exp(-\lambda y)$	$(0,\infty)$	k/λ	k/λ^2
Normal	$(2\pi\sigma^2)^{-1/2}\exp(-(y-\mu)^2/2\sigma^2)$	$(-\infty,\infty)$	μ	σ^2

Other Ways to Characterize a Random Variable

Notation	Name	Relationship
$f(t)$	density	$\frac{d}{dt}F(t) = h(t)S(t)$
$F(t)$	distribution	$\int_0^t f(u)du$
$S(t)$	survivor	$1 - F(t)$
$h(t)$	hazard	$f(t)/S(t)$
$H(t)$	integrated hazard	$\int_0^t h(u)du = -\ln S(t)$

The Ohio State University

RT Distributions

Ex-Gaussian

$$f(t \mid \mu, \sigma^2, \tau) = \frac{1}{2\tau} \exp\left((2\mu + \sigma^2/\tau - 2t)/2\tau\right) \text{erfc}\left(\frac{\mu + \sigma^2/\tau - t}{\sqrt{2}\sigma}\right) \mathcal{I}\left(t \in \mathbb{R}\right)$$

Inverse normal (Wald)

$$f(t \mid a, \nu) = \left(\frac{a^2}{2\pi t^3}\right)^{1/2} \exp\left[\frac{-a^2(t - (a/\nu))^2}{2(a/\nu)^2 t}\right] \mathcal{I}\left(t > 0\right)$$

Drift diffusion model (without interesting variance)

$$\begin{aligned} &f(t, 0 \mid a, z, \nu, T_{er}) = \\ &\quad \frac{\pi}{a^2} \exp(-z\nu) \sum_{k=1}^{\infty} k \sin\left(\frac{\pi z k}{a}\right) \exp\left(-\frac{1}{2}(\nu^2 + \pi^2 k^2/a^2)\right) (t - T_{er}) \mathcal{I}\left(t > T_{er}\right) \end{aligned}$$

Linear Ballistic Accumulator

$$\begin{aligned} &f(t \mid A, b, \nu, s) = \\ &\quad \frac{1}{A}\left[-\nu\Phi\left(\frac{b - A - t\nu}{ts}\right) + s\phi\left(\frac{b - A - t\nu}{ts}\right) + \nu\Phi\left(\frac{b - t\nu}{ts}\right) - s\phi\left(\frac{b - t\nu}{ts}\right)\right] \mathcal{I}\left(t > 0\right) \end{aligned}$$

THE OHIO STATE UNIVERSITY

Exercises

- Choose an RT density function and program it in R
- Simulate your favorite RT model and save the data for later.

Things to think about:

- What is the *support* of the random variable you are simulating?
- What values are the parameters of the distribution permitted to take on? How are you going to make sure the values are appropriate?

Outline

Wagenmakers et al. (2004)

- 6 participants
- 8 digit stimuli: (1,2,3,4) and (6,7,8,9)
- 3 tasks (1024 observations per task)
  1. Simple RT: press a key when a digit is presented
  2. Choice RT: press right key for even numbers, left for odd
  3. Temporal estimation: press a key 1s after digit onset
- Response stimulus interval (RSI)
  1. Short: $U[550, 950]$
  2. Long: $U[1150, 1550]$

Outline

Cumulative Distribution Function

$F(t) = Pr(T \leq t)$
The easiest way:

$$\hat{F}(t) = \frac{1}{N} \sum_{i=1}^{N} \mathcal{I}\left(T_i \leq t\right)$$

Slightly harder:

$$\hat{F}(t \mid h) = \int_{u=0}^{t} \hat{f}(t \mid h)$$

Exercise: Estimate the CDF of the data you simulated. Is it accurate? How would you know?

Probability Density Function

The easy but un-useful way (histogram):

$$\hat{f}(t \mid \mathbf{c}) = \frac{1}{N(c_j - c_{j-1})} \sum_{i=1}^{N} \mathcal{I}\left(c_{j-1} < T_i \leq c_j\right) \text{ for } c_{j-1} < t < c_j$$

The less easy but very useful way:

$$\hat{f}(t \mid h) = \frac{1}{Nh} \sum_{i=1}^{N} K\left(\frac{t - T_i}{h}\right)$$

The function K is a *kernel* that integrates to 1.
Exercise: Estimate the PDF (density) of the data you simulated. Check its accuracy using the function you coded.

An Important Side Note

Given an *iid* sample of data $T = \{T_1, T_2, \ldots, T_N\}$, the *likelihood* is

$$\mathcal{L}(\theta \mid T) = \prod_{i=1}^{n} f(T_i \mid \theta),$$

where f is the PDF of T_i and θ is a vector of the distribution's parameters.
Exercise: Compute the likelihood of the sample of data you simulated, using a PDF of your choice. Things to think about:

- What's the best way to compute the product of a large set of tiny numbers?
- What are you going to do if the density of one of your observations is very very small?

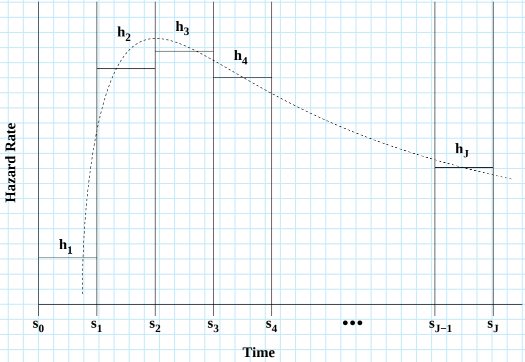
The Ohio State University

The Hazard Function

Kernel estimate:

$$\hat{h}(t \mid h) = \frac{\hat{f}(t \mid h)}{1 - \hat{F}(t \mid h)}$$

Houpt et al. 2016: Piecewise Exponential Model



Semi-parametric Model

$$h(t \mid \theta) = \frac{f(t \mid \theta)}{1 - F(t \mid \theta)} \Rightarrow f(t \mid \theta) = h(t \mid \theta)\,[1 - F(t \mid \theta)]$$

Semi-parametric Model

$$h(t \mid \theta) = \frac{f(t \mid \theta)}{1 - F(t \mid \theta)} \Rightarrow f(t \mid \theta) = h(t \mid \theta)\left[1 - F(t \mid \theta)\right]$$

$$\mathcal{L}(\theta \mid t_1, t_2, \ldots, t_n) = \prod_{i=1}^{n} h(t_i \mid \theta)\left[1 - F(t_i \mid \theta)\right]$$

Semi-parametric Model

$$h(t \mid \theta) = \frac{f(t \mid \theta)}{1 - F(t \mid \theta)} \Rightarrow f(t \mid \theta) = h(t \mid \theta)\left[1 - F(t \mid \theta)\right]$$

$$\begin{aligned}
& L(h \mid t_1, t_2, \ldots, t_n) \\
&= \prod_{i=1}^{n} f(t_i \mid h) \\
&= \prod_{i=1}^{n} \prod_{j=1}^{J} \left[h_j \exp\left\{ - \sum_{n=1}^{j-1} h_n (s_n - s_{n-1}) - h_j (t_i - s_{j-1}) \right\} \right]^{I(s_{j-1} < t_i < s_j)}
\end{aligned}$$

for defined bin boundaries $s_0, s_1, \ldots, s_J$ and hazard rates $h_1, \ldots, h_J$.

Hazard Estimates

50 Samples per distribution