Contents

Chapter 1

Introduction and Asymptotic Analysis for Tukey Depth in the Bivariate Exponential Distribution

1.1 Introduction

In this chapter, we study the Tukey (half-space) depth for a bivariate distribution with independent exponential marginals. Let

$$X \sim \mathrm{Exp}(\lambda_1) \quad \text{and} \quad Y \sim \mathrm{Exp}(\lambda_2),$$

with joint density

$$f_{X,Y}(x,y) = \lambda_1 \lambda_2 \, e^{-\lambda_1 x - \lambda_2 y}, \quad x, y \geq 0.$$

The Tukey depth of a point is defined as the minimum probability mass of any closed half-plane that contains the point. In our approach, we consider a family of lines passing through the point (a, b) and study the probability mass below the line. (Note that in general the depth is given by

$$D(a,b) = \min\{P, 1 - P\},$$

but here we focus on the computation of one side's probability as a function of the line parameters.)

1.2 Definition of the Probability Mass $I(k)$

Consider a line through (a, b) given by

$$y - b = k\,(x - a), \tag{1.1}$$

where $k = \tan\theta$. In the first quadrant, the region below this line forms a triangle. The x-intercept is found by setting $y = 0$:

$$0 - b = k\,(x - a) \quad \Longrightarrow \quad x = a - \frac{b}{k}.$$

We define

$$x_0 = a - \frac{b}{k}, \quad \text{(assumed positive)}.$$

Then, for each $x \in [0, x_0]$, the corresponding y runs from 0 up to

$$y_{\max}(x) = k(x - a) + b.$$

Thus, the probability mass below the line is given by

$$I(k) = \int_{x=0}^{x_0} \int_{y=0}^{k(x-a)+b} \lambda_1 \lambda_2 \, e^{-\lambda_1 x - \lambda_2 y} \, dy \, dx. \tag{1.2}$$

Evaluating the inner integral,

$$\int_0^{k(x-a)+b} \lambda_2 e^{-\lambda_2 y} \, dy = 1 - e^{-\lambda_2 \left[k(x-a)+b\right]},$$

we obtain

$$I(k) = \int_{x=0}^{x_0} \lambda_1 e^{-\lambda_1 x} \left[1 - e^{-\lambda_2 \left(k(x-a)+b\right)}\right] dx. \tag{1.3}$$

1.3 Differentiation via the Leibniz Rule

Since the upper limit $x_0 = a - \frac{b}{k}$ depends on k, we differentiate $I(k)$ using the Leibniz rule. In general, if

$$I(k) = \int_0^{x_0(k)} f(x, k) \, dx,$$

then

$$\frac{dI}{dk} = \int_0^{x_0(k)} \frac{\partial f(x, k)}{\partial k} \, dx + f\big(x_0(k), k\big) \, \frac{dx_0}{dk}.$$

In our case,

$$f(x, k) = \lambda_1 e^{-\lambda_1 x} \left[1 - e^{-\lambda_2 \left(k(x-a)+b\right)}\right].$$

The k-dependence is entirely in the exponential term. Differentiating,

$$\frac{\partial}{\partial k} \left[1 - e^{-\lambda_2 \left(k(x-a)+b\right)}\right] = \lambda_2 (x - a) \, e^{-\lambda_2 \left(k(x-a)+b\right)}.$$

Thus,

$$\frac{\partial f(x, k)}{\partial k} = \lambda_1 e^{-\lambda_1 x} \, \lambda_2 (x - a) \, e^{-\lambda_2 \left(k(x-a)+b\right)}.$$

Also, differentiating $x_0 = a - \frac{b}{k}$ with respect to k gives

$$\frac{dx_0}{dk} = \frac{b}{k^2}.$$

At $x = x_0$,

$$k(x_0 - a) + b = k\Big(a - \frac{b}{k} - a\Big) + b = -b + b = 0,$$

so that

$$1 - e^{-\lambda_2 \left(k(x_0-a)+b\right)} = 0.$$

Hence, the boundary term vanishes and we obtain

$$\frac{dI}{dk} = \int_{x=0}^{x_0} \lambda_1 e^{-\lambda_1 x} \, \lambda_2 (x - a) \, e^{-\lambda_2 \left(k(x-a)+b\right)} \, dx. \tag{1.4}$$

Setting $\frac{dI}{dk} = 0$ yields an implicit equation for the optimal k^*.

1.4 Asymptotic Analysis: Two Regimes

We now discuss how the behavior of the function

$$g(x) = k(x-a) + b$$

affects the optimal condition for k in two regimes.

1.4.1 Small-Scale Regime

In the regime where the values of $g(x)$ remain small for all $x \in [0, x_0]$, we can expand the exponential in a Taylor series:

$$e^{-\lambda_2\, g(x)} \approx 1 - \lambda_2\, g(x).$$

Then,

$$1 - e^{-\lambda_2\, g(x)} \approx \lambda_2\, g(x) = \lambda_2\big[k(x-a) + b\big].$$

Substitute this into (1.3):

$$I(k) \approx \int_0^{x_0} \lambda_1 e^{-\lambda_1 x}\, \lambda_2\big[k(x-a) + b\big]\, dx.$$

Differentiating this approximate form with respect to k and setting the derivative equal to zero leads (after some algebra) to the condition

$$k^* \approx \frac{\lambda_1}{\lambda_2}.$$

Since $k = \tan\theta$, this implies

$$\tan\theta^* \approx \frac{\lambda_1}{\lambda_2}.$$

1.4.2 Large-Scale Regime

When the values of $g(x) = k(x-a)+b$ become large for most x in $[0, x_0]$ except near the upper limit $x = x_0$ (where $g(x_0) = 0$), the exponential $e^{-\lambda_2\, g(x)}$ decays rapidly except in a narrow region near x_0. To focus on this region, we define a new variable

$$t = x_0 - x,$$

so that when $x = x_0$, $t = 0$ and when $x = 0$, $t = x_0$. In terms of t, we have

$$x = x_0 - t.$$

Then,

$$g(x) = k\big((x_0 - t) - a\big) + b = k(x_0 - a) + b - k\,t.$$

Since $k(x_0 - a) + b = 0$, it follows that

$$g(x) \approx -k\,t.$$

(Here, the negative sign indicates that $g(x)$ decreases linearly to 0 at $t = 0$; we are interested in the magnitude.) Hence, the exponential factor becomes

$$e^{-\lambda_2\, g(x)} \approx e^{\lambda_2 k\, t}.$$

Because $e^{\lambda_2 k\,t}$ grows rapidly as t increases, and our original integrals in (1.3) involve $1-e^{-\lambda_2\,g(x)}$, the dominant contribution arises from values of t very close to 0 (i.e. x near x_0), where the approximation $g(x) \approx k\,t$ holds. Using Laplace's method, we approximate the integral by evaluating the slowly varying factor $\lambda_1 e^{-\lambda_1 x}$ at $x = x_0$ and integrating the exponential term over t:

$$\int_{t=0}^{\epsilon} e^{\lambda_2 k\,t}\,dt \approx \frac{e^{\lambda_2 k\,\epsilon}-1}{\lambda_2 k},$$

for a small ϵ. In the limit where the decay is very rapid (large $\lambda_2 k$), the effective support is very near $t = 0$. Differentiating the logarithm of the dominant contribution with respect to k yields a correction term of order $1/k$. Expressing this condition in terms of θ (since $k = \tan\theta$) leads to an optimality condition that includes a logarithmic correction. (A detailed derivation shows that the condition takes the form)

$$\tan\theta^* \approx \frac{\lambda_1}{\lambda_2}\left(1+\Delta\right),$$

with

$$\Delta \sim \frac{1}{K}\ln\frac{\lambda_1}{\lambda_2},$$

where K represents the characteristic scale of $g(x)$ in the region near x_0. In a fully detailed derivation, this correction is found to be of the form

$$\tan\theta^* \approx \frac{\lambda_1}{\lambda_2}\left(1+\frac{1}{c}\ln\frac{\lambda_1}{\lambda_2}\right),$$

if one were to reintroduce the geometric scale c. However, since we are not introducing a separate projection variable, the key takeaway is that in the large-scale regime the optimal k^* deviates from λ_1/λ_2 by a logarithmic correction that arises from the steep decay of the exponential term in a small neighborhood near $x = x_0$.

1.5 Summary

To summarize, working solely in terms of k, a, and b:

- The integration region is given by $x \in [0, x_0]$ with $x_0 = a - \frac{b}{k}$ and, for each x, $y \in [0,\, k(x-a)+b]$.
- The probability mass below the line is

$$I(k) = \int_{x=0}^{x_0} \lambda_1 e^{-\lambda_1 x}\left[1-e^{-\lambda_2\left(k(x-a)+b\right)}\right]dx.$$

- Differentiating $I(k)$ with respect to k using the Leibniz rule yields an implicit equation for the optimal k^*.
- In the regime where $k(x-a)+b$ remains small (small-scale regime), a Taylor expansion shows that $k^* \approx \frac{\lambda_1}{\lambda_2}$.
- In the regime where $k(x-a)+b$ is large except near $x = x_0$ (large-scale regime), by substituting $t = x_0 - x$ and applying Laplace's method we find that the integral is dominated by a narrow region near $t = 0$. Differentiating the logarithm of this dominant contribution introduces a logarithmic correction in the optimal condition.

While the exact form of the logarithmic correction requires a full derivation, the key point is that in the large-scale regime the optimal slope k^* deviates from λ_1/λ_2 by a term proportional to a logarithm of λ_1/λ_2 (scaled by the characteristic size of the integration region near x_0). This completes our extended explanation entirely in terms of k, a, and b.

Chapter 2

Projection of an Exponential Point and Hypoexponential Distribution

In this chapter we consider the same joint density

$$f_{X_1,X_2}(x_1,x_2) = \lambda_1\lambda_2\, e^{-\lambda_1 x_1 - \lambda_2 x_2}, \quad x_1, x_2 \geq 0,$$

and analyze the depth of a fixed point (a,b) in R^2. Here, the random vector is denoted by

$$\mathbf{X} = (X_1, X_2),$$

and we study the projection of $\mathbf{X}$ onto an arbitrary direction.

2.1 Projection onto a Given Direction and Half-Space Definition

Let $u = (\cos\theta, \sin\theta)$ be an arbitrary unit vector. The projection of the random vector $\mathbf{X} = (X_1, X_2)$ onto u is given by

$$Z = u^T\mathbf{X} = X_1\cos\theta + X_2\sin\theta.$$

For a fixed point (a,b), its projection is

$$u^T(a,b) = a\cos\theta + b\sin\theta.$$

Define the half-space associated with the direction u as

$$H(u) = \{x \in R^2 : \, u^T x \geq u^T(a,b)\}.$$

Then the Tukey depth of (a,b) is defined as

$$D(a,b) = \inf_{u\in S^1}\Big\{P\big(x \in R^2 : \, u^T x \geq u^T(a,b)\big)\Big\} = \inf_{u\in S^1}\Big\{1 - F\big(u^T(a,b)\big)\Big\},$$

where F is the cumulative distribution function (CDF) of the projection Z.

2.2 Projection and the Hypoexponential Distribution

Since X_1 and X_2 are independent exponential random variables, the projection

$$Z = X_1 \cos\theta + X_2 \sin\theta$$

can be viewed as a sum of two independent scaled exponentials. Define

$$\mu_1 = \frac{\lambda_1}{\cos\theta} \quad \text{and} \quad \mu_2 = \frac{\lambda_2}{\sin\theta}.$$

If $\mu_1 \neq \mu_2$, then Z follows a hypoexponential distribution with CDF

$$F(z) = 1 - \frac{\mu_2}{\mu_2 - \mu_1}\, e^{-\mu_1 z} + \frac{\mu_1}{\mu_2 - \mu_1}\, e^{-\mu_2 z}, \quad z \geq 0.$$

In the special case when $\mu_1 = \mu_2 = \mu$ (which occurs if $\lambda_1 \sin\theta = \lambda_2 \cos\theta$), the CDF becomes

$$F(z) = 1 - (1 + \mu z)e^{-\mu z}.$$

Thus, the probability that $\mathbf{X}$ falls in the half-space is

$$P\{u^T\mathbf{X} \geq u^T(a,b)\} = 1 - F\big(u^T(a,b)\big),$$

and the depth is

$$D(a,b) = \inf_{u\in S^1} \Big\{1 - F\big(a\cos\theta + b\sin\theta\big)\Big\}.$$

2.3 Summary and Optimization Approach

The key idea is that the projection $Z = X_1 \cos\theta + X_2 \sin\theta$ follows a hypoexponential distribution (with the special case $\mu_1 = \mu_2$ considered separately). Consequently, the probability mass in the half-space determined by u is

$$1 - F\big(a\cos\theta + b\sin\theta\big),$$

and the Tukey depth of (a,b) is obtained by taking the infimum of this expression over all directions u (or equivalently, over θ). In practice, one may evaluate

$$1 - F\big(a\cos\theta + b\sin\theta\big)$$

explicitly using the hypoexponential CDF formulas provided above and then use numerical methods (such as grid search or optimization routines) to find

$$D(a,b) = \inf_{\theta\in[0,2\pi)} \Big\{1 - F\big(a\cos\theta + b\sin\theta\big)\Big\}.$$

This formulation reduces the multidimensional depth problem to a one-dimensional optimization problem involving the hypoexponential CDF.