

# Races or Tournaments: Identification & Estimation

Andrea Blasco ([ablasco@fas.harvard.edu](mailto:ablasco@fas.harvard.edu))

March 27, 2017

# Contents

|          |   |          |
|----------|---|----------|
| <b>1</b> | <b>Identification</b>                           | <b>3</b> |
| 1.1      | Example with the Uniform distribution . . . . . | 5        |
| 1.2      | Example with Beta distribution . . . . .        | 6        |
| <b>2</b> | <b>Estimation</b>                               | <b>6</b> |
| 2.1      | More examples . . . . .                         | 7        |

# 1 Identification

In this section, we discuss the identification problem, which consists in the ability to use the observed data to identify one or more of the parameters of the contest model presented in Section ?? . In particular, we determine conditions necessary to identify parameters of the cost function and ability distribution (the “primitives” of the model) and how these vary across different competition modes: race, tournament, and tournament with a minimum-entry requirement.

Our observations come from the results of the experimental rooms of the challenge described in Section ?? . For each room ( $l = 1, \dots, L$ ), we observe the number  $Y_l$  of participating competitors among  $N_l$  registered competitors who have been assigned to some treatment  $\tau_l$ . The results are reported in Table 1.

| Room name | Treatment  | Registrants | Participants |
|-----------|------------|-------------|--------------|
| Group 1A  | Race       | 9           | 2            |
| Group 1B  | Race       | 10          | 2            |
| Group 1C  | Race       | 10          | 5            |
| Group 1D  | Race       | 10          | 1            |
| Group 1E  | Race       | 15          | 4            |
| Group 1F  | Race       | 15          | 3            |
| Group 1G  | Race       | 15          | 4            |
| Group 1H  | Race       | 15          | 5            |
| Group 2A  | Tournament | 10          | 3            |
| Group 2B  | Tournament | 10          | 5            |
| Group 2C  | Tournament | 10          | 5            |
| Group 2D  | Tournament | 10          | 2            |
| Group 2E  | Tournament | 15          | 5            |
| Group 2F  | Tournament | 15          | 4            |
| Group 2G  | Tournament | 15          | 4            |
| Group 2H  | Tournament | 15          | 5            |
| Group 3A  | Reserve    | 10          | 1            |
| Group 3B  | Reserve    | 10          | 3            |
| Group 3C  | Reserve    | 10          | 3            |
| Group 3D  | Reserve    | 10          | 2            |
| Group 3E  | Reserve    | 15          | 4            |
| Group 3F  | Reserve    | 15          | 6            |
| Group 3G  | Reserve    | 15          | 3            |
| Group 3H  | Reserve    | 15          | 5            |

Table 1: Participating competitors by room and treatment group. Registered competitors signed up to the challenge and were randomly assigned to non-overlapping rooms. Each room was then randomly assigned to a treatment. A participant is a registered competitor who has made at least one submission during the eight-day submission period of the challenge.

We assume the number of participating competitors  $Y_l$  at a room  $x_l = c(\tau_l, z_l)$ , where  $\tau_l$  is the treatment randomly assigned to the room and  $z_l$  is a vector of additional room characteristics, is distributed as a binomial variable with parameters  $N_l$  and  $p(x_l, \theta)$  where  $p(x, \theta)$  describes the probability for a competitor to participate in the challenge at a room described by  $x$ . The model is then the following:

$$Y_l \sim \text{Binomial}(N_l, p(x_l, \theta)) \quad \text{for each } l = 1, \dots, L. \quad (1)$$

In binomial linear models, the probability function  $p(x, \theta)$  is related to a linear prediction through a one-to-one transformation called the link function (see xxxx). In the MS theoretical approach, our preliminary

analysis of the model showed that the predictor is not generally linear and it is modeled as follows:

$$p(x_l, \theta) = 1 - F(\hat{a}_l; \theta) \quad (2)$$

where  $F$  is a parametric distribution function of the latent ability variable (privately observed by each player) with  $\theta$  being a vector of distributional parameters and  $\hat{a}_l$  being the marginal type of the room.

In each room, the marginal type is determined by the following zero profit condition:

$$R(\hat{a}) = c_a(\hat{a})c_y(\underline{y})c_t(\bar{t}) \quad (3)$$

where the  $R(\hat{a})$  the expected revenue of the marginal type must be equal to the costs evaluated at the deadline and the target, if any.

In a contest with two prizes, the expected revenue is

$$R(\hat{a}) = \alpha F_{(n-1:n-1)}(\hat{a}) + (1 - \alpha)[1 - F_{(n-1:n-1)}(\hat{a})]F_{(n-2:n-2)}(\hat{a}). \quad (4)$$

The left hand side of the above equation describes the value of winning the first prize  $\alpha$  times the probability of the marginal player  $\hat{a}$  being the winner. The second term is the value of winning the second prize  $1 - \alpha$  times the probability of the marginal player being the second ranked competitor.

Because abilities are drawn at random and independently, the distribution of the order statistics on the left hand side depends on the distribution  $F$  and, hence, it is known up to the parameters  $\theta$  that we want to estimate from the data.

Also, the cost function of the marginal type is determined by the room characteristics  $x$  (e.g., the problem, whether there's a deadline or not, whether's there's a minimum entry requirement or not). The following linear relationship is assumed:

$$C(\underline{y}, \bar{t}) = \beta x_l. \quad (5)$$

That is, the costs is linear in the the room characteristics  $x_l$  with the parameters  $\beta$  that we also want to estimate from the data.

If we divide all sides by the ability cost  $c_a$  the left hand side of (3), then we use the following definition:

$$G(\hat{a}, F_\theta) \equiv c_a(\hat{a})^{-1}R(\hat{a}) \quad (6)$$

where the function  $G(\cdot, F_\theta)$  is strictly monotone in the marginal type, as we showed in the discussion of the model. Because monotone in its first argument, the function is also invertible. This leads to the following

$$\hat{a} = G^{-1}(\beta x_l, F_\theta). \quad (7)$$

Using (7) with (2) leads to:

$$p(x_l, \omega) = 1 - F_\theta(G^{-1}(\beta x_l, F_\theta)) \quad (8)$$

with  $\omega = (\theta, \beta)$ .

## 1.1 Example with the Uniform distribution

Suppose  $\alpha = 1$  to simplify computations. Imagine that the distribution  $F_\theta$  is uniform on  $(0, \theta)$  and the cost ability  $c_a(x) = 1/x$ . Then the zero profit condition (3) becomes

$$\left(\frac{\hat{a}}{\theta}\right)^{n-1} \hat{a} = \beta x_l \quad (9)$$

Solving for the marginal type gives:

$$\hat{a} = (\beta x_l \theta^{n-1})^{1/n} \quad (10)$$

So, the probability of participation (2) becomes

$$p(x_l, \theta) = 1 - F_\theta(\hat{a}) \quad (11)$$

$$= 1 - \frac{(\beta x_l \theta^{n-1})^{1/n}}{\theta} \quad (12)$$

$$= 1 - \left(\frac{\beta x_l}{\theta}\right)^{1/n}. \quad (13)$$

Because it is non-linear in its parameters, the model is not identifiable [need proof]. However, we can re-parametrize the model with  $\beta/\theta \equiv \eta$  to make the parameter  $\eta$  identifiable.

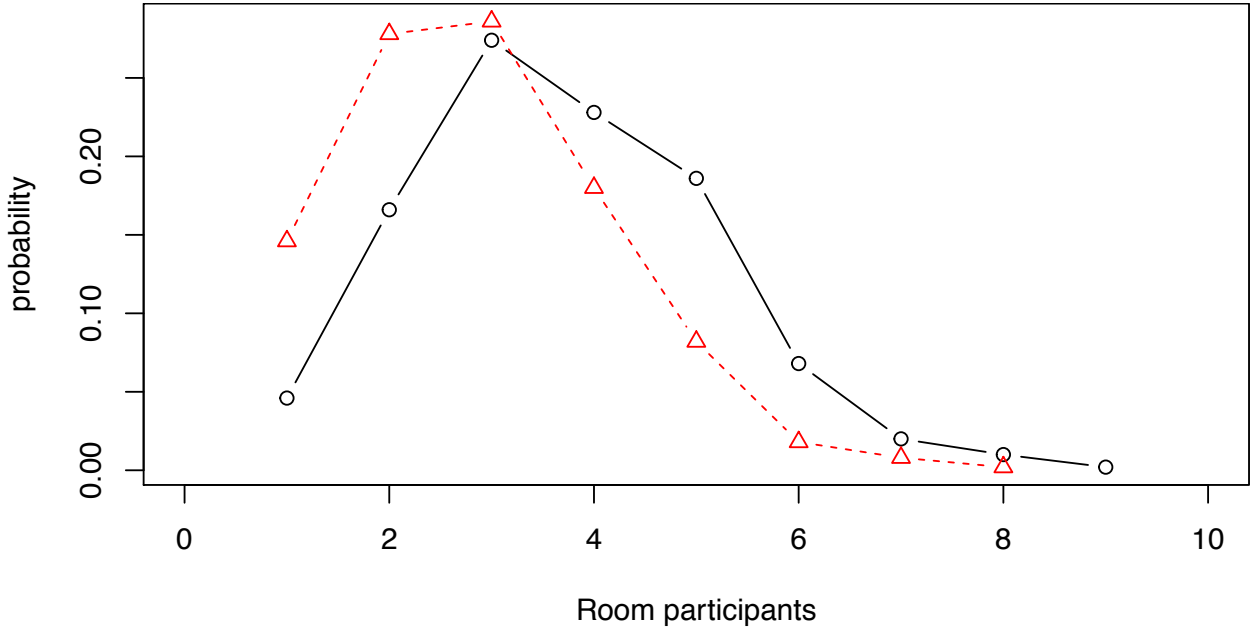


Figure 1: Simulated distribution of participants with uniform distribution. Parameters are  $\eta = 0.1$  and  $n = 15$ . Data simulated 500 times. Dashed curve had higher costs ( $x_l = 1.5$ ) than the solid curve ( $x_l = 0.5$ ).

## 1.2 Example with Beta distribution

In this example, we assume abilities are drawn from a lognormal distribution with parameters  $\mu$  and  $\sigma$ .

## 2 Estimation

Our aim is to estimate the parameters  $\theta$  and  $\beta$  and to evaluate whether costs are different between the three competition modes under study: race, tournament, tournament + requirement.

It is natural to estimate the parameters  $\theta$  by maximum likelihood because the ability distribution (and therefore the distribution of the outcomes) is known. The estimation criterion used here is the maximization of the *deviance*.

The deviance is:

$$D(\theta) = -2 \sum Y \log\left(\frac{Np(x, \theta)}{Y}\right) + (N - Y) \log\left(\frac{N - Np}{N - Y}\right). \quad (14)$$

Note that the deviance is a function of the likelihood (see xxx pg. xxx).

```
# Define likelihood
structreg <- function(x, y, wt = rep(1, length(y)), intercept = TRUE, start = rep(0, p), .
  fmin <- function(beta, X, y, w) {
    p <- 1 - (beta %*% X)^(1/n) # Function of parameters
    -sum(2 * w * ifelse(y, log(p), log(1-p)))
  }
# gmin <- function(beta, X, y, w) {
#   # Gradient here!
# }
if(is.null(dim(x))) dim(x) <- c(length(x), 1)
dn <- dimnames(x)[[2]]
if(!length(dn)) dn <- paste("Var", 1:ncol(x), sep="")
p <- ncol(x) + intercept
if(intercept) {x <- cbind(1, x); dn <- c("(Intercept)", dn)}
if(is.factor(y)) y <- (unclass(y) != 1)
# fit <- optim(start, fmin, gmin, X = x, y = y, w = wt, method = "BFGS", ...)
fit <- optim(0.5, fmin, X = x, y = y, w = wt, method = "BFGS", ...)
names(fit$par) <- dn
cat("\nCoefficients:\n"); print(fit$par)
cat("\nResidual Deviance:", format(fit$value), "\n")
cat("\nConvergence message:", fit$convergence, "\n")
invisible(fit)
}

# Create data
ilogit <- function(x) exp(x) / (1+exp(x))
n <- 50000
```

```

competitors <- 5
x1 <- rchisq(competitors, df=3)
x2 <- runif(n)
x3 <- rnorm(n)
X <- cbind(rep(1, length(x1)), x1, x2, x3)
betas <- c(-2.5, 0.1, 0.25, 0.5)
fc <- ilogit(X %*% betas) ## This the fixed cost
p <- 1 - fc^(1/competitors)
y <- rbinom(n=n, size=competitors, prob=p)

# Objective function
fmin <- function(beta, X, y, w, competitors) {
  p <- 1 - ilogit(X %*% beta)^(1/competitors) # Function of parameters
  -sum(2 * w * ifelse(y, log(p), log(1-p)))
}

b.start <- runif(length(betas))
(out <- optim(b.start, fmin, X=X, y=y, competitors=competitors, w=rep(1, length(y)))$par)

## [1] -12.8162255 0.4280835 1.0236303 2.0644232

rbind(out/competitors, betas)

##           [,1]      [,2]      [,3]      [,4]
##      -2.563245 0.0856167 0.2047261 0.4128846
## betas -2.500000 0.1000000 0.2500000 0.5000000

```

## 2.1 More examples

Suppose  $F$  is lognormal( $\mu, \sigma$ ), then the zero profit condition is:

$$\left[ \frac{1}{2} + \frac{1}{2} \left( \frac{\log \hat{a} - \mu}{\sqrt{2}\sigma} \right) \right] = a^{-1/(n-1)} K_l \quad (15)$$

```

# xxx
zeroprofit <- function(x, mu, sigma, n) {
  x * plnorm(x, meanlog=mu, sdlog=sigma)^(n-1)
}
for (i in 3:10)
  curve(zeroprofit(x, i, 10, 10), add=TRUE, lty=i)

```