

Регрессионные модели с ограничениями и регуляризацией при больших наборах регрессоров. Задача восстановления состава инвестиционного портфеля.

Рыбка Елизавета

Московский Физико-Технический Институт

Выпускная квалификационная работа на соискание степени бакалавра.

Научный руководитель: д.т.н, профессор Моттль В.В.

28 июня 2018 г.

Задача восстановления состава инвестиционного портфеля

Активы — ценные бумаги (акции, облигации, векселя и пр.), $i = 1, \dots, n$, торгуемые на биржевом рынке и характеризующиеся в каждый момент времени своей биржевой ценой.

Доходность на некотором периоде владения t :

$$x_{t,i} = \frac{z_{t,i}^{end} - z_{t,i}^{beg}}{z_{t,i}^{beg}}, \text{ относительная скорость изменения цены}$$

Активов на биржевом рынке n — тысячи.

Инвестиционный портфель — долевое распределение капитала некоторой компании среди части ценных бумаг $\hat{n} \ll n$
(как правило, глубокая тайна)

$$\beta = (\beta_1, \dots, \beta_n) \in \mathbb{R}^n; \quad \beta_i \geq 0, \quad \sum_{i=1}^n \beta_i = 1$$

Задача — угадать распределение капитала по совокупности наблюдений за доходностями активов и портфеля в целом.

Returns Based Style Analysis (Уильям Шарп)

Returns Based Style Analysis по Шарпу

Линейная модель Шарпа: $y \cong \sum_{i=1}^n \beta_i x_i$

Совокупность наблюдений: $(y_t, x_{t,i}, i = 1, \dots, n), t = 1, \dots, T$
— доходности портфеля, активов.

Задача квадратичного программирования по Шарпу (RBSA)
в предположении $n < T$:

$$\begin{cases} (\hat{\beta}_1, \dots, \hat{\beta}_n) = \arg \min \sum_{t=1}^T (y_t - \sum_{i=1}^n \beta_i x_{t,i})^2 \\ \sum_{i=1}^n \beta_i = 1 \\ \beta_i \geq 0, i = 1, \dots, n \end{cases} \quad (1)$$

Однако в реальности $n \gg T$ (тысячи против сотен).

В этой ситуации невозможно найти малое подмножество активов $\hat{n} \ll n$, в действительности составляющих портфель.

Необходимо учитывать априорные предположения о том, как администратор мог бы составить инвестиционный портфель.

В магистерских диссертациях Ильи Пугача и Алексея Морозова, эта проблема была названа Factor Search: $|\hat{\mathbb{I}}| = \hat{n} \ll n = |\mathbb{I}|$

Основная идея — при оценивании долевого распределения капитала учитывать статистическую информацию о доходности исследуемого портфеля и доходностях всех биржевых активов.

$$\begin{cases} \beta^T \Sigma \beta - \mu \bar{x}^T \beta + c (\mathbf{y} - \mathbf{X}^T \beta)^T (\mathbf{y} - \mathbf{X}^T \beta) \rightarrow \min(\beta) \\ \mathbf{1}^T \beta = 1; \quad \beta \geq \mathbf{0} \end{cases} \quad (2)$$

$$\mathbf{y} = (y_1, \dots, y_T) \in \mathbb{R}^T \quad \mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_T) \in \mathbb{R}^{n \times T}$$

Здесь Σ ($n \times n$) и \bar{x} (n) — результат технического анализа; ковариационная матрица и вектор математических ожиданий колебаний доходностей активов в прошлом;

Идея данной работы — дополнить технический анализ прошлых доходностей активов экспертным мнением о предпочтительном скрытом составе портфеля.

Математическая формализация экспертных суждений об ожидаемом скрытом составе портфеля

Предлагается — дополнить оценки ковариационной матрицы и среднего вектора доходностей активов (Σ, \bar{x}) матрицей и вектором той же структуры $(\mathbf{B}; \mathbf{z})$, формализующих мнение эксперта о разумном составе портфеля.

Пусть эксперт выразил свое мнение относительно:

Волатильности
доходностей активов

доход-

Несовместности пар ак-
тивов

Предпочтительном
составе портфеля

$$\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix} \in \mathbb{R}^n$$

$$\mathbf{R} = \begin{pmatrix} 1 & r_{12} & \cdots & r_{1n} \\ & 1 & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix}$$

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

Агрегированная матрица: $\mathbf{B} = [\text{Diag}(\mathbf{d})]^T \mathbf{R} [\text{Diag}(\mathbf{d})]$

Математическая формализация экспертных суждений об ожидаемом скрытом составе портфеля

Агрегированная матрица: $\mathbf{B} = [\text{Diag}(\mathbf{d})]^T \mathbf{R} [\text{Diag}(\mathbf{d})]$

Структура этой матрицы аналогична структуре ковариационной матрицы:

$$\mathbf{B} = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} = \begin{pmatrix} d_1 d_1 r_{11} & \dots & d_1 d_n r_{1n} \\ \vdots & \ddots & \vdots \\ d_n d_1 r_{n1} & \dots & d_n d_n r_{nn} \end{pmatrix} \quad (3)$$

$$\text{Сравним: } \Sigma = \begin{pmatrix} \sigma_1 \sigma_1 \rho_{11} & \dots & \sigma_1 \sigma_n \rho_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_n \sigma_1 \rho_{n1} & \dots & \sigma_n \sigma_n \rho_{nn} \end{pmatrix} \quad (4)$$

Математическая формализация экспертных суждений об ожидаемом скрытом составе портфеля

Пусть эксперт выразил свое мнение относительно:

Волатильности
ностей активов

доход-

Несовместности пар ак-
тивов

Предпочтительном
составе портфеля

$$\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix} \in \mathbb{R}^n$$

$$\mathbf{R} = \begin{pmatrix} 1 & r_{12} & \cdots & r_{1n} \\ & 1 & \cdots & r_{2n} \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix}$$

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

Агрегированная матрица: $\mathbf{B} = [\text{Diag}(\mathbf{d})]^T \mathbf{R} [\text{Diag}(\mathbf{d})]$

Фактически эти матрицы определяют суждение эксперта о скрытом портфеле: $0 \leq \mu < \infty$ степень недоверия к доходностям активов

$$\begin{cases} \beta^T \mathbf{B} \beta - \mu \mathbf{z}^T \beta \rightarrow \min(\beta) \\ \mathbf{1}^T \beta = 1; \quad \beta \geq \mathbf{0} \end{cases} \quad (5)$$

Пример выбора матрицы В

Всего биржевых активов $\mathbb{I} = \{i = 1, \dots, n\}$, $n = 650$

Предположим, что эксперт разбил это множество на $m = 15$ групп, объединяющих активы, сходные по некоторым социально-деловым представлениям: $\mathbb{I} = \mathbb{I}_1 \cup \dots \cup \mathbb{I}_m$, $\mathbb{I}_k \cap \mathbb{I}_l = \emptyset$, $k \neq l$.

Эксперт предполагает, что в портфель вряд ли входят более одного представителя каждой группы.

В терминах матрицы \mathbf{R} несовместимости пар активов это означает, что штраф $r_{i,j}$ для i, j из одной группы должен быть очень большим, например 0.9999, поскольку $r_{i,i} = 1$

Соответственно, диагональные (а) и недиагональные (б) блоки \mathbf{R} имеют вид:

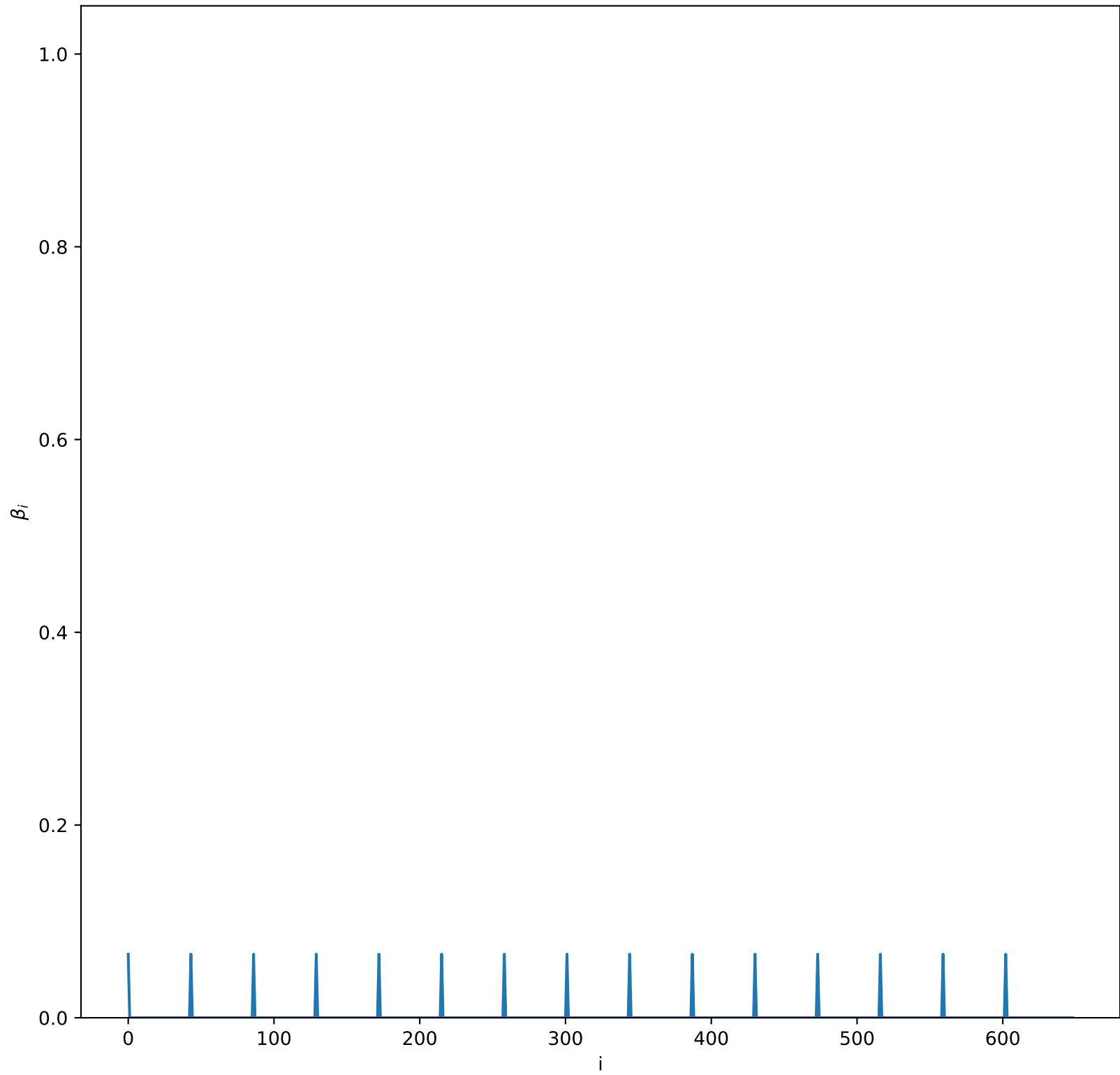
1	0.99	...	0.99
0.99	1	...	0.99
\vdots	\vdots	\ddots	\vdots
0.99	0.99	...	1

(а)

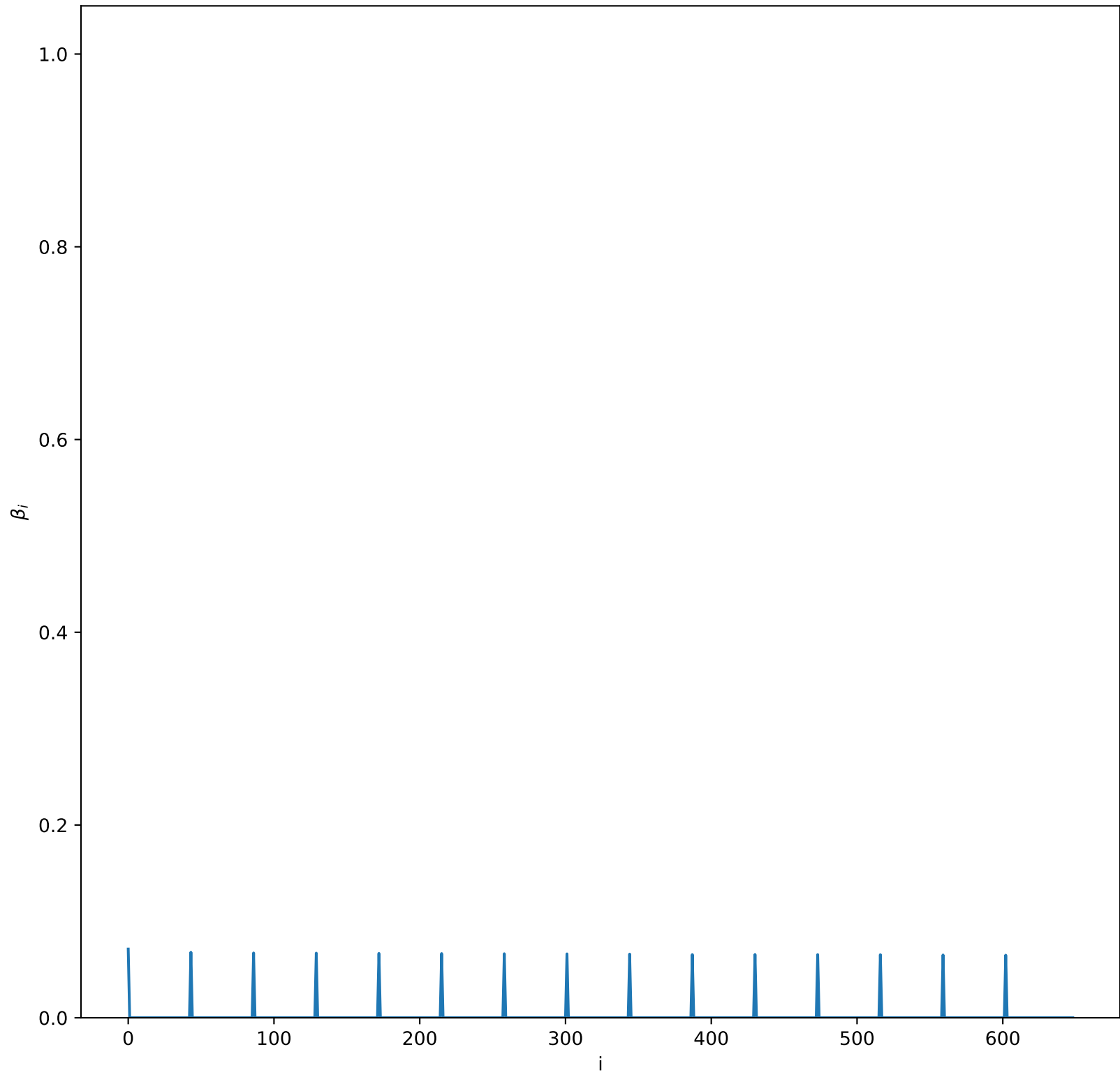
0	0	...	0
0	0	...	0
\vdots	\vdots	\ddots	\vdots
0	0	...	0

(б)

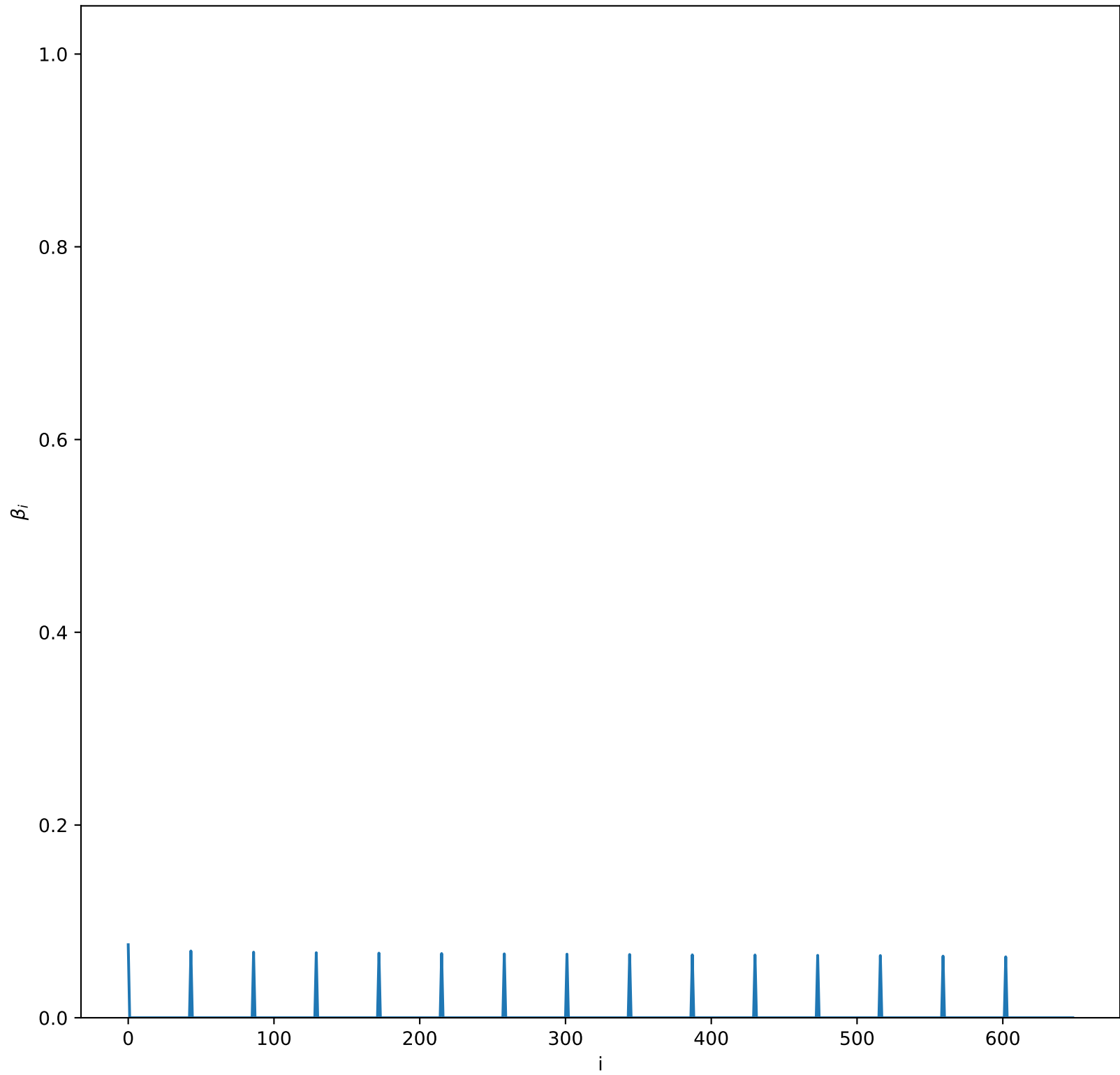
$\mu = 0.00$



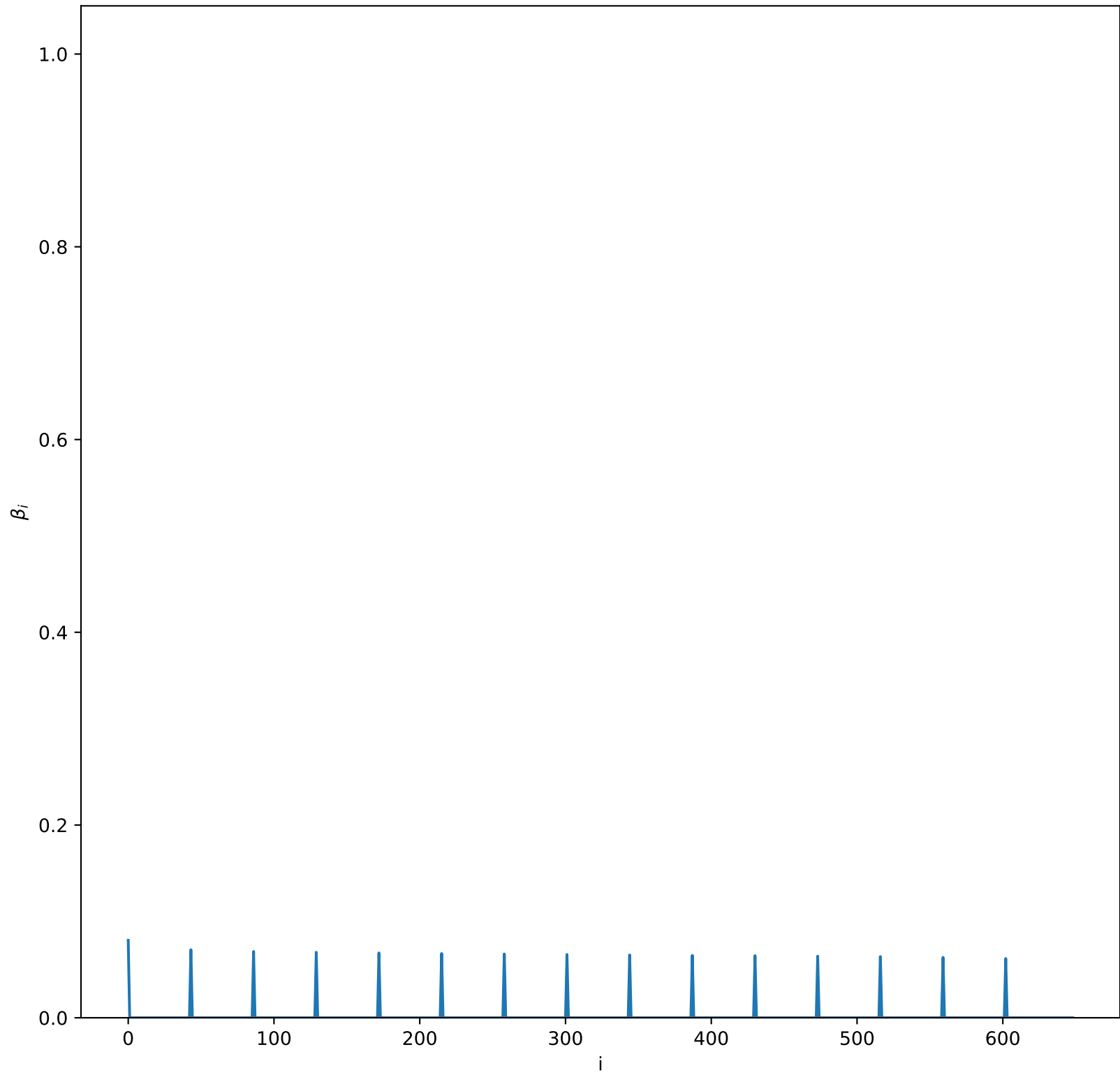
$\mu = 0.01$



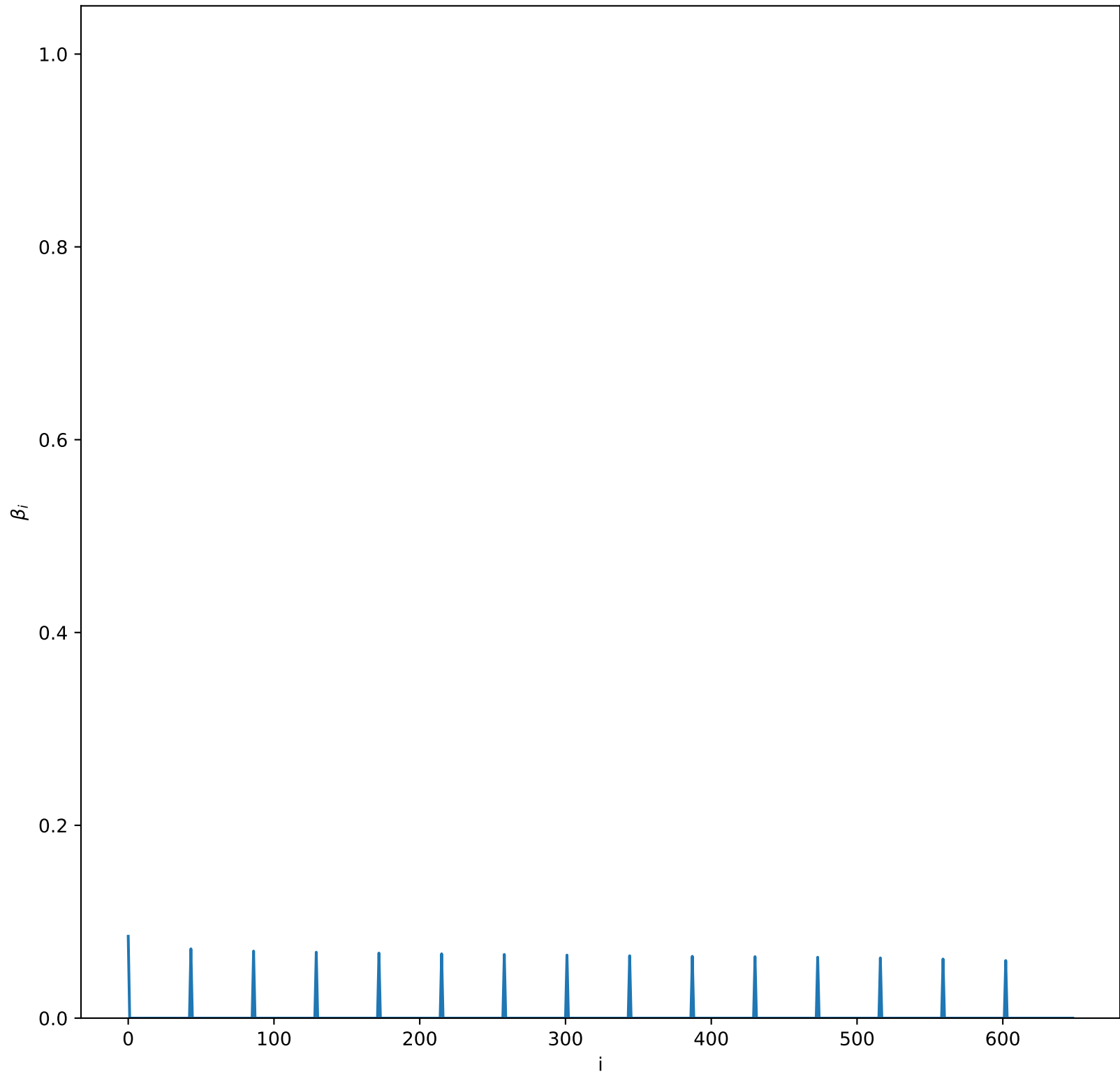
$\mu = 0.02$



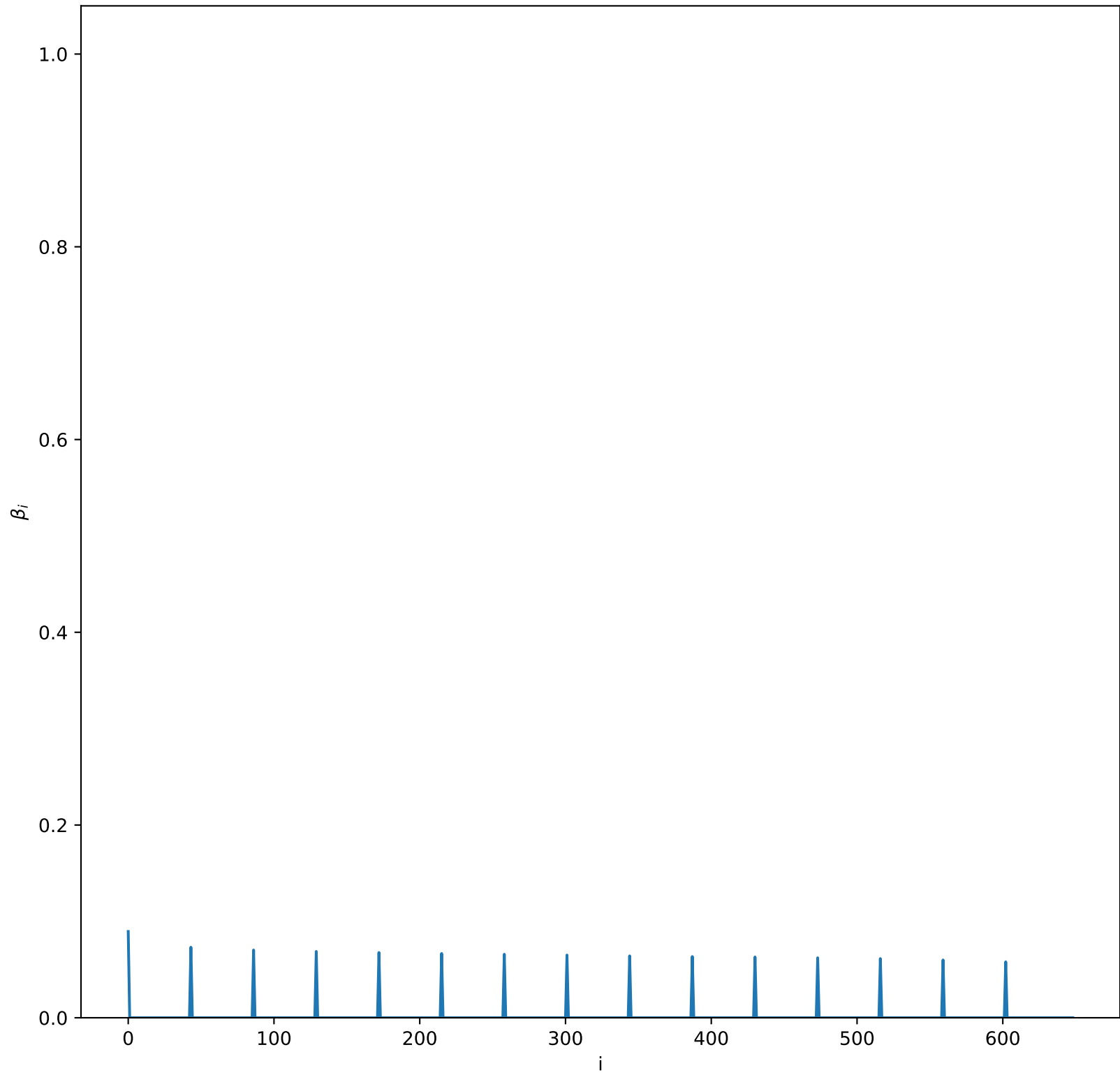
$\mu = 0.03$



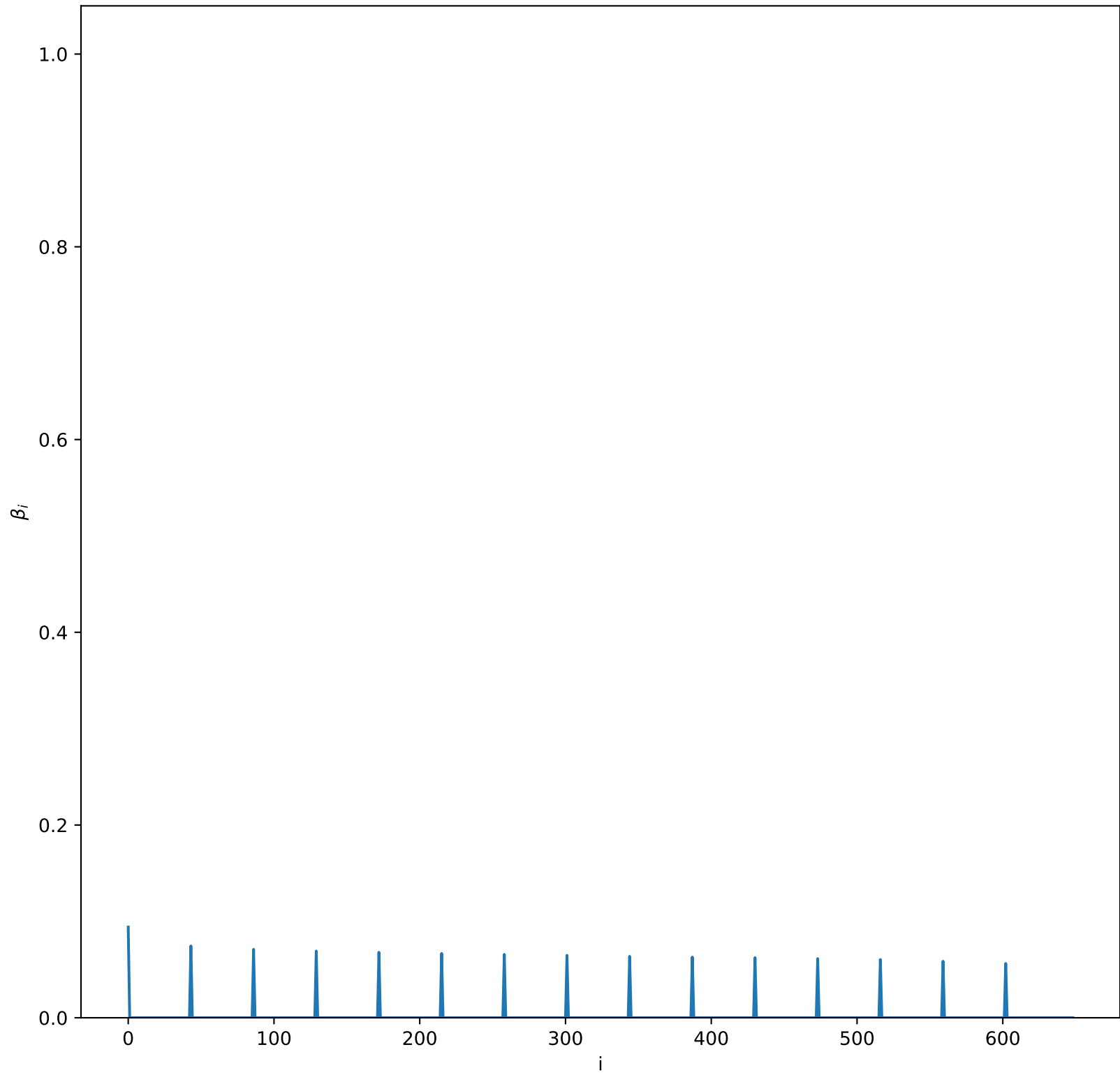
$\mu = 0.04$



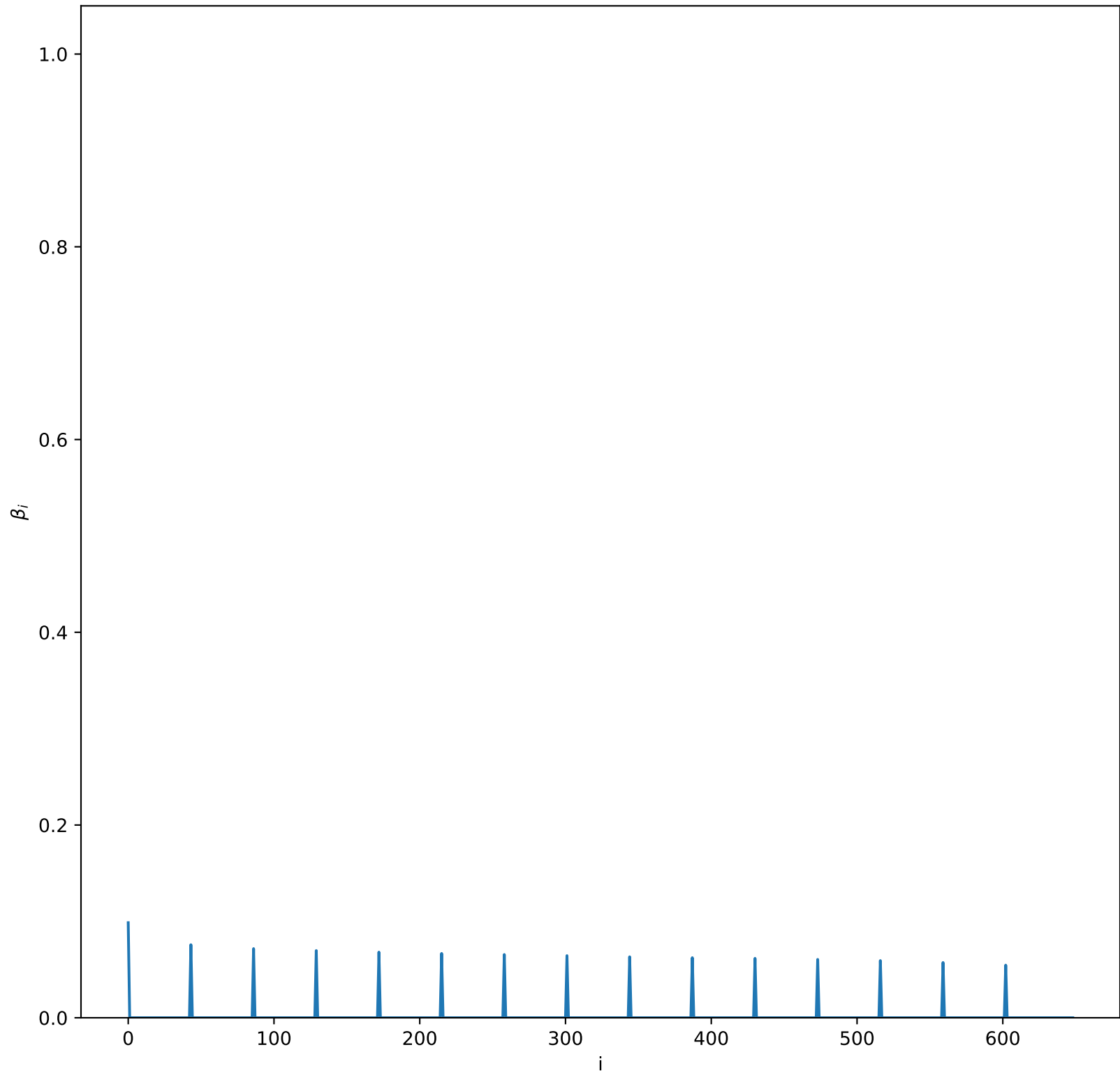
$\mu = 0.05$



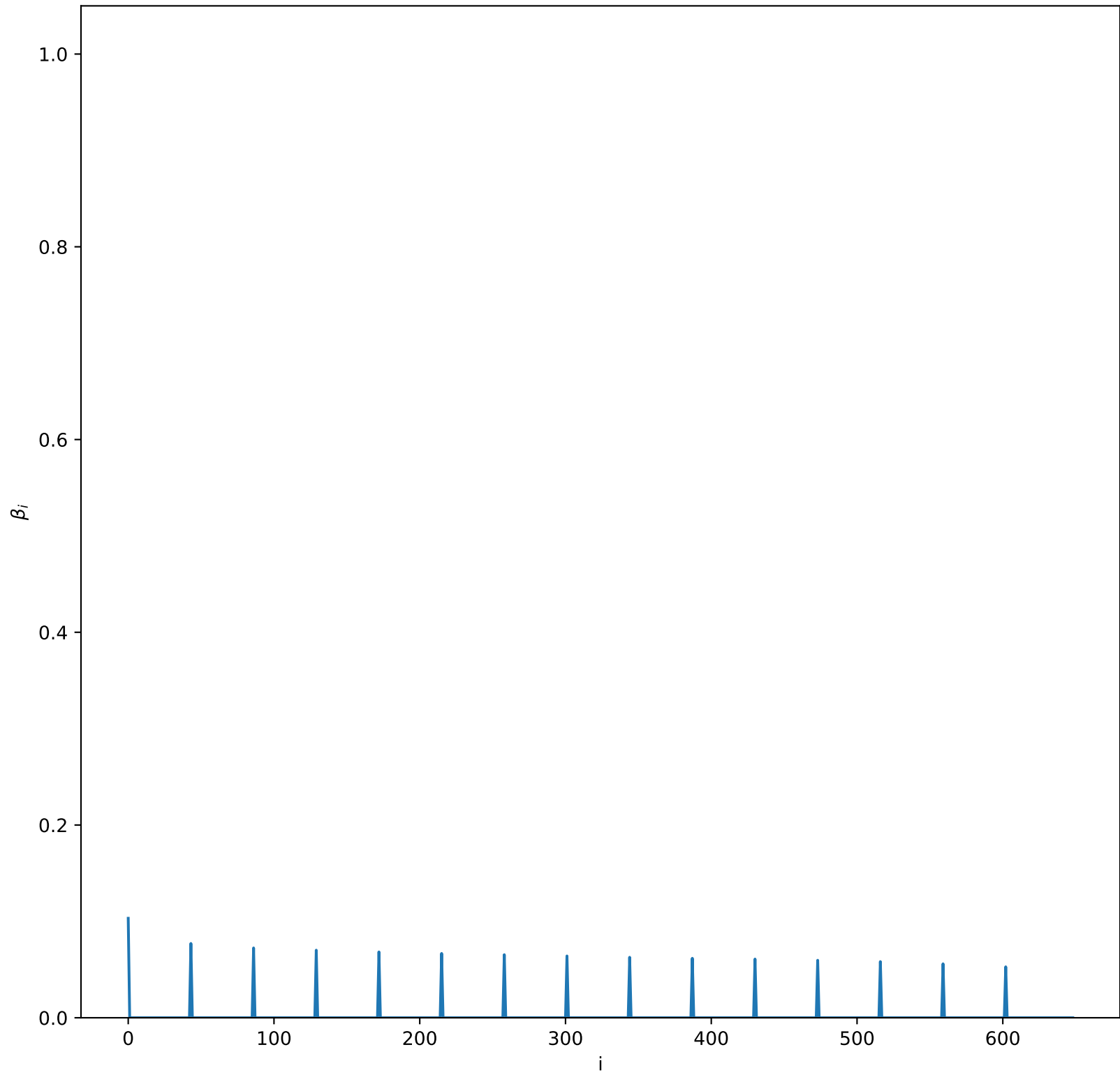
$\mu = 0.06$



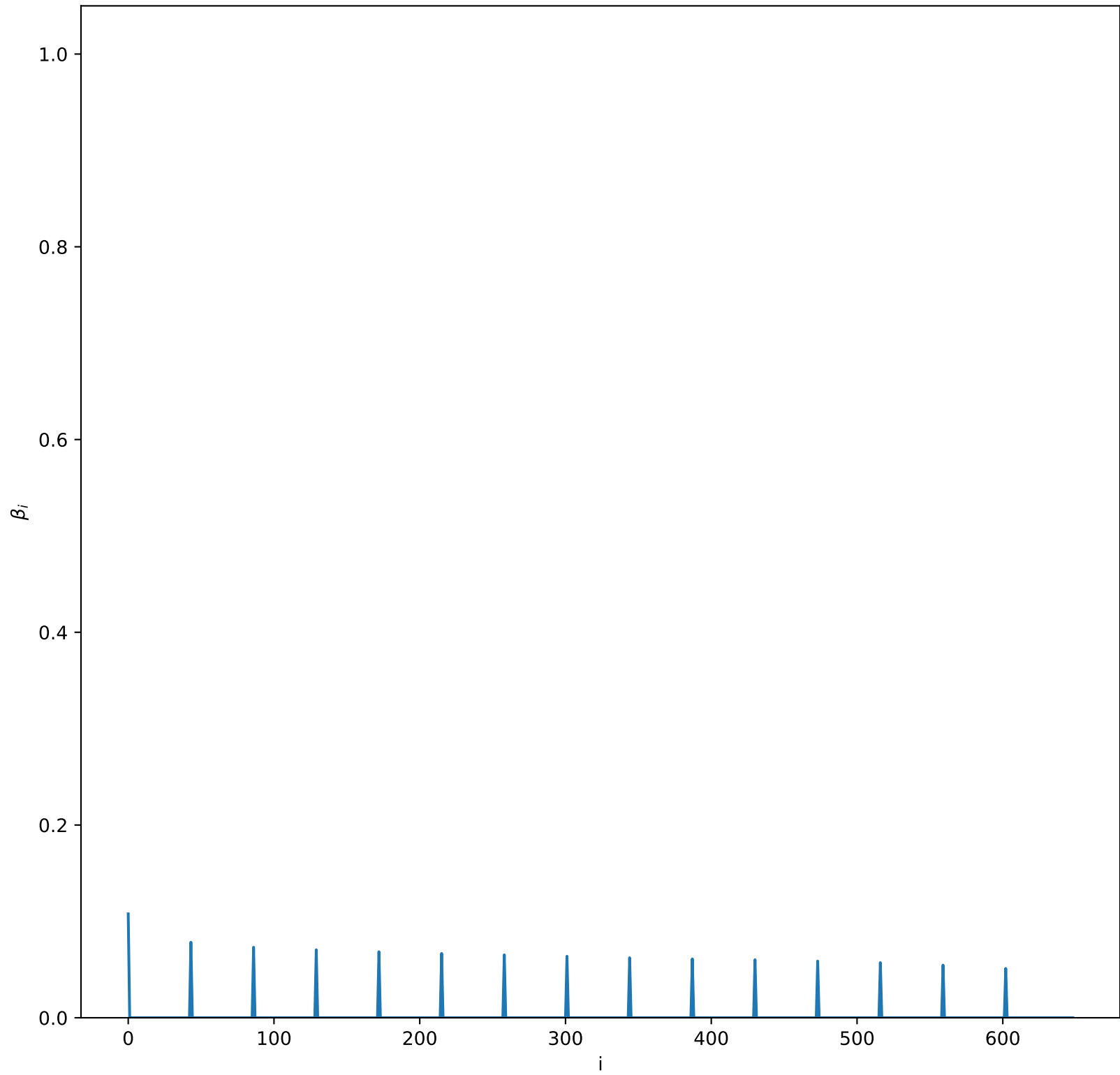
$\mu = 0.07$



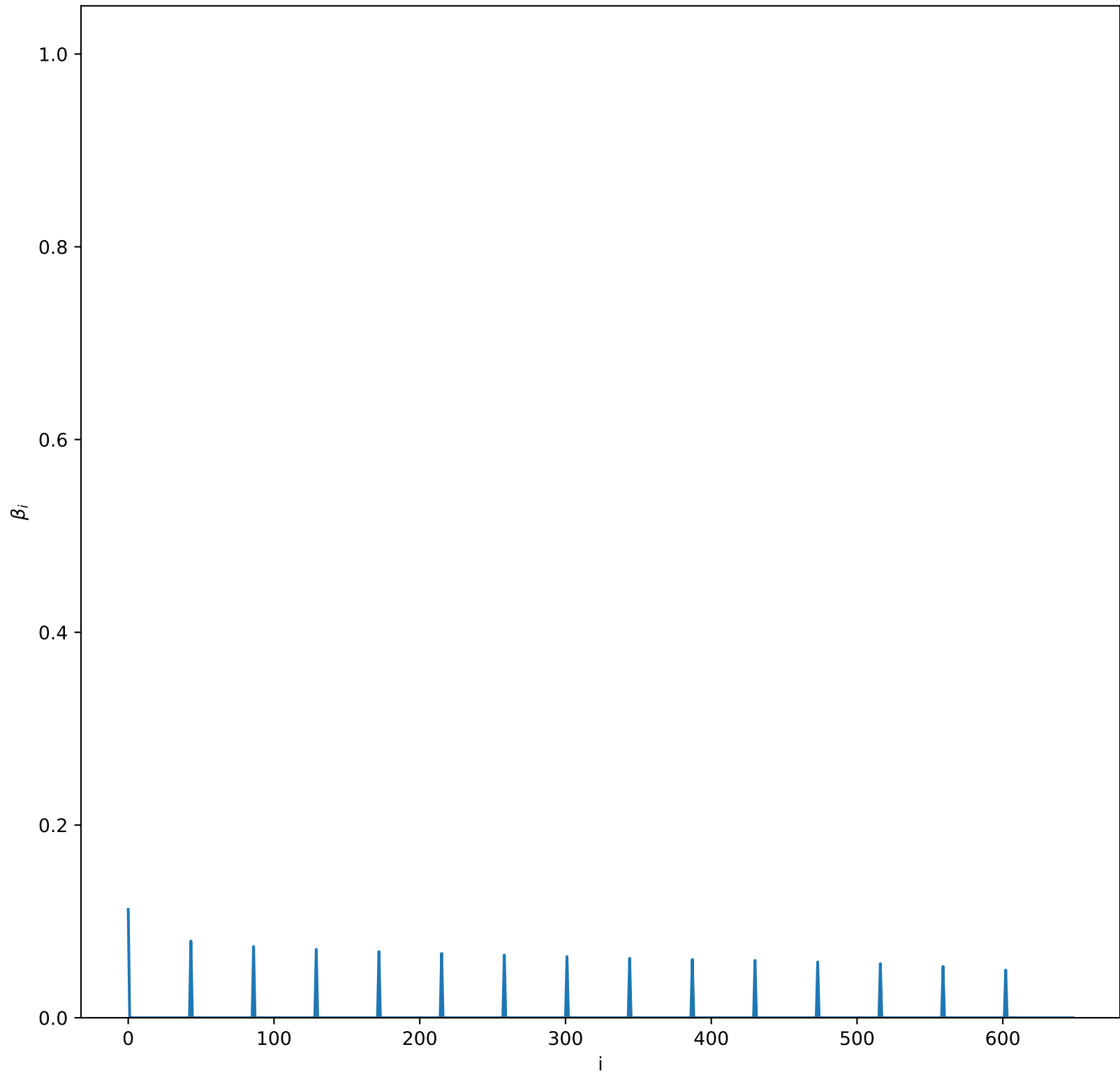
$\mu = 0.08$



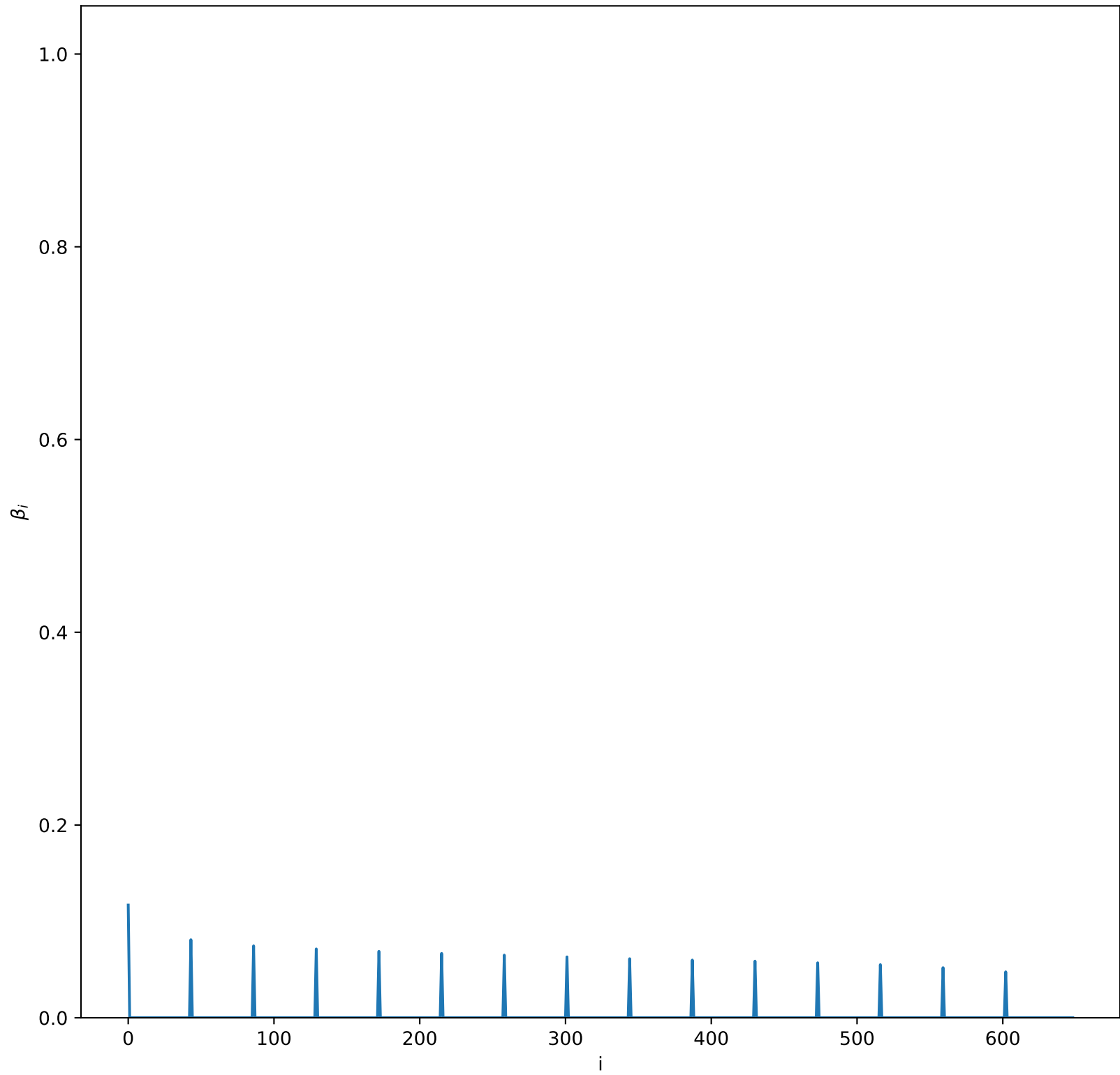
$\mu = 0.09$



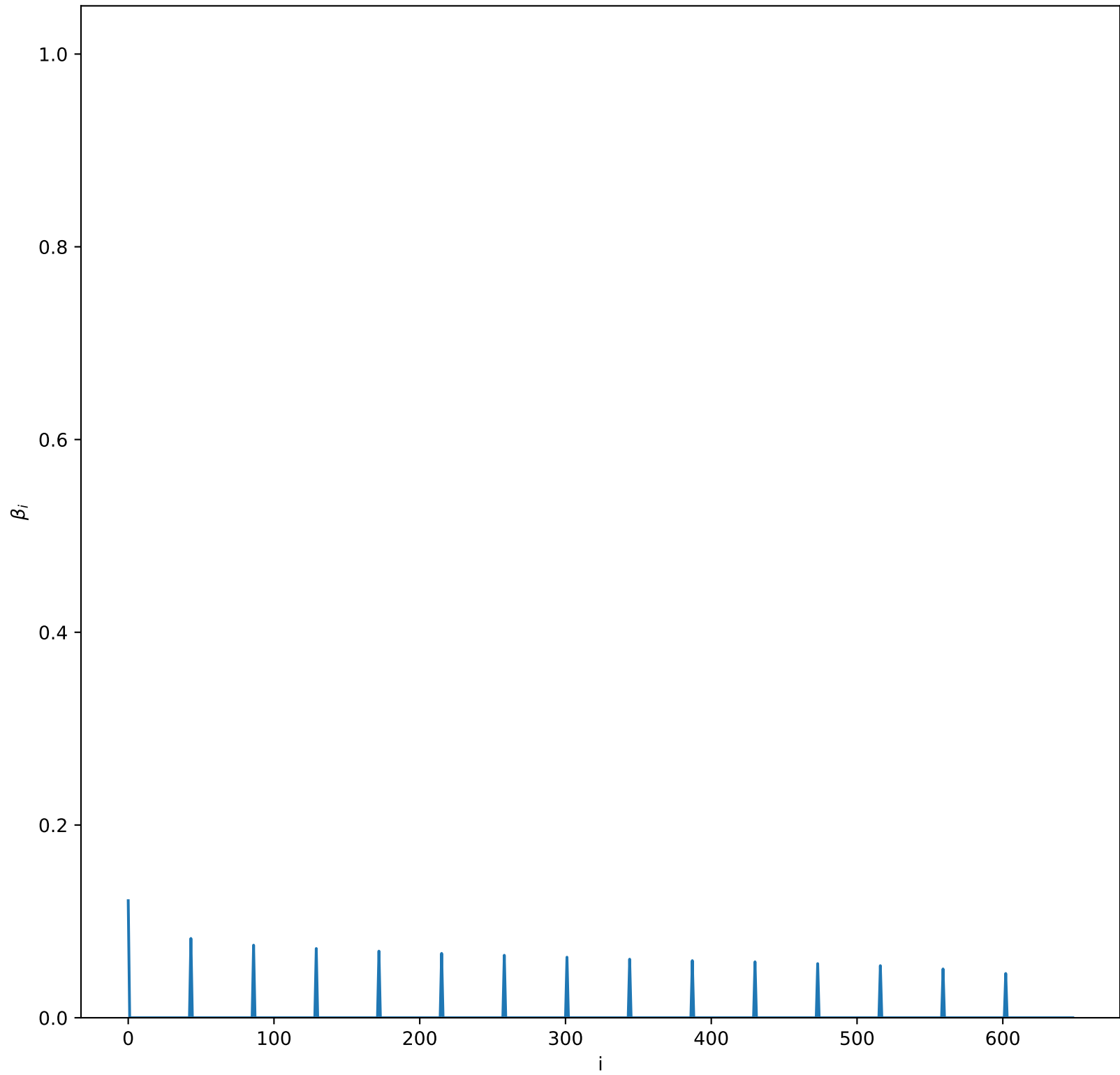
$\mu = 0.10$



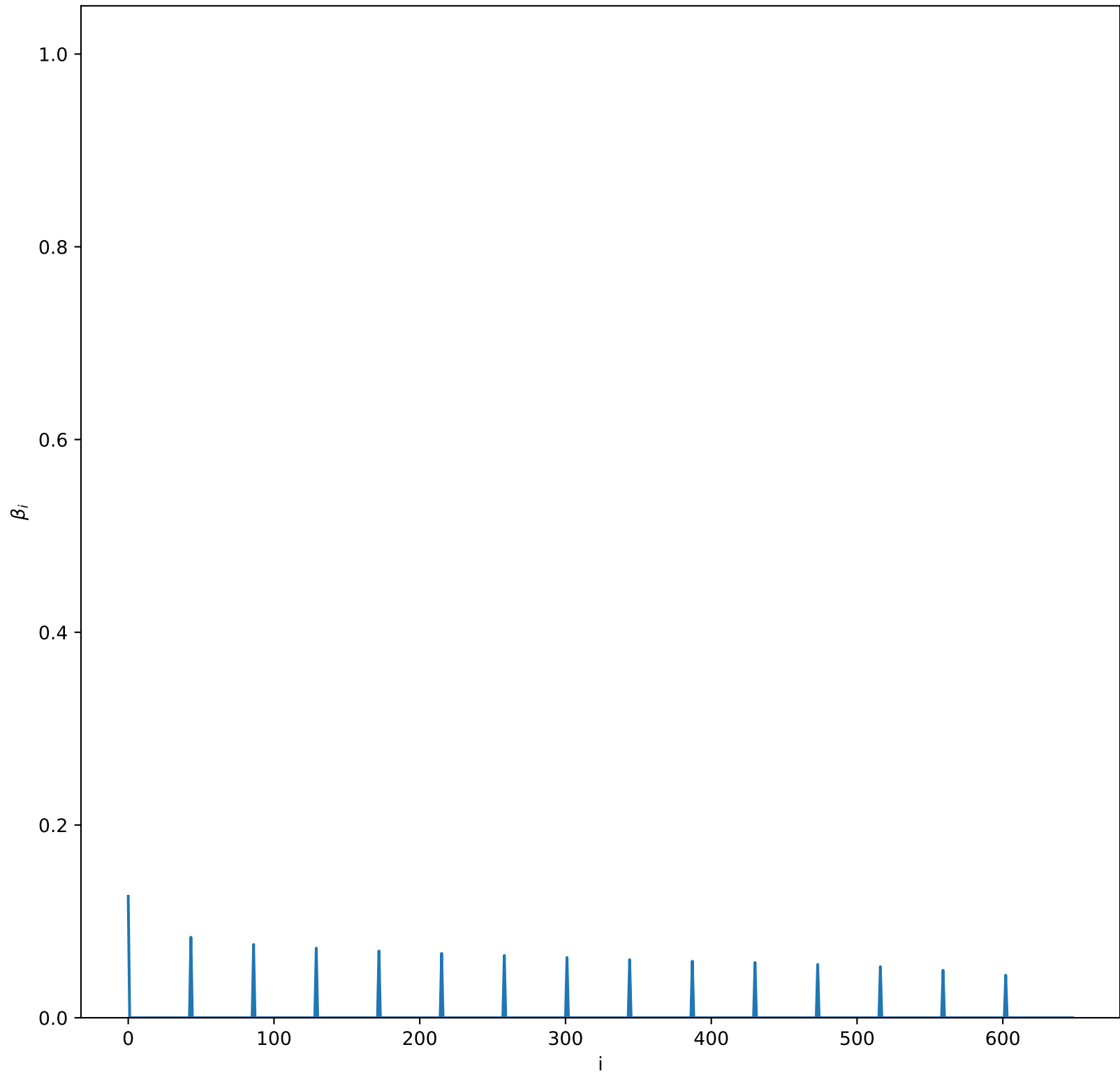
$\mu = 0.11$



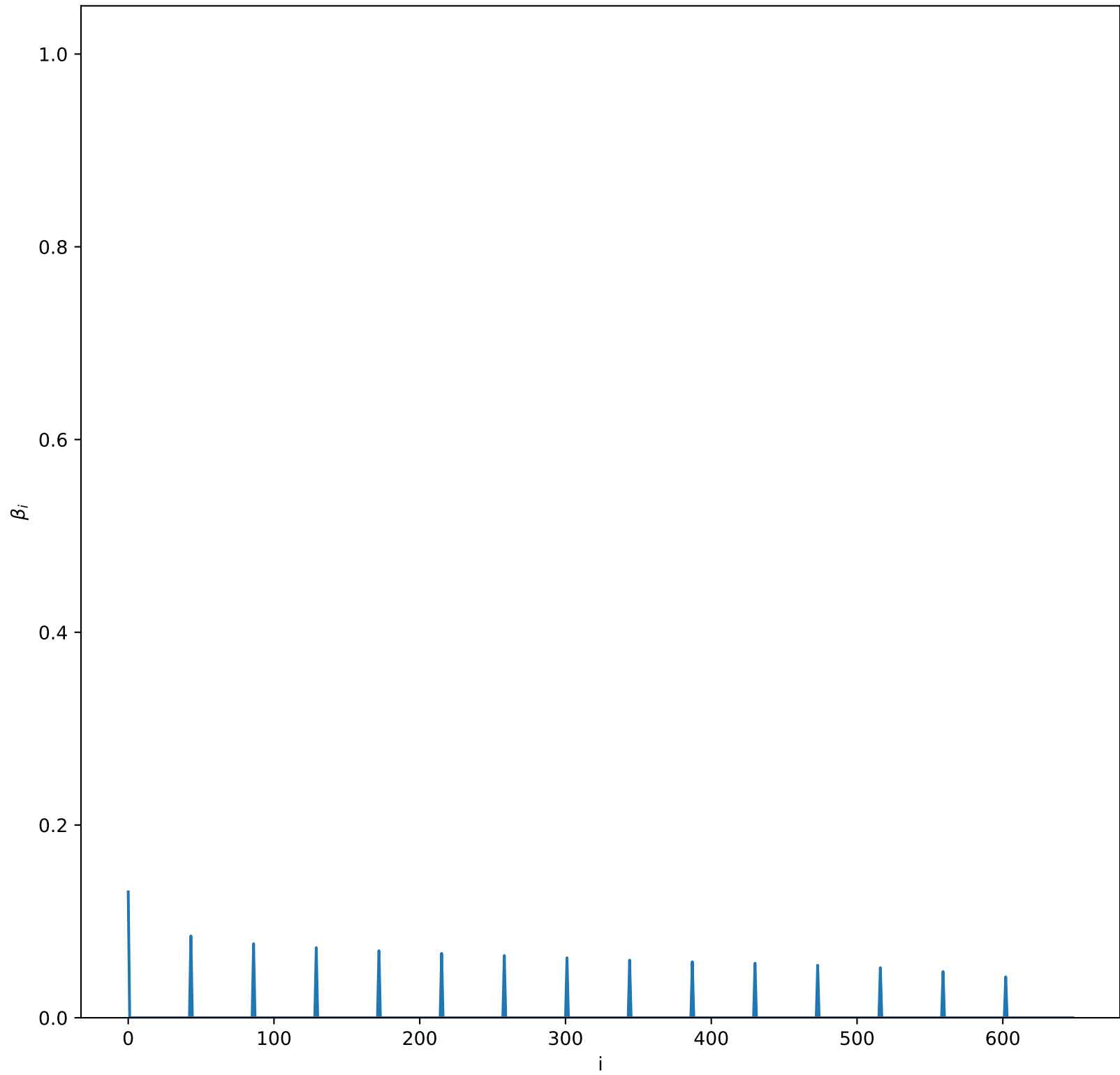
$\mu = 0.12$



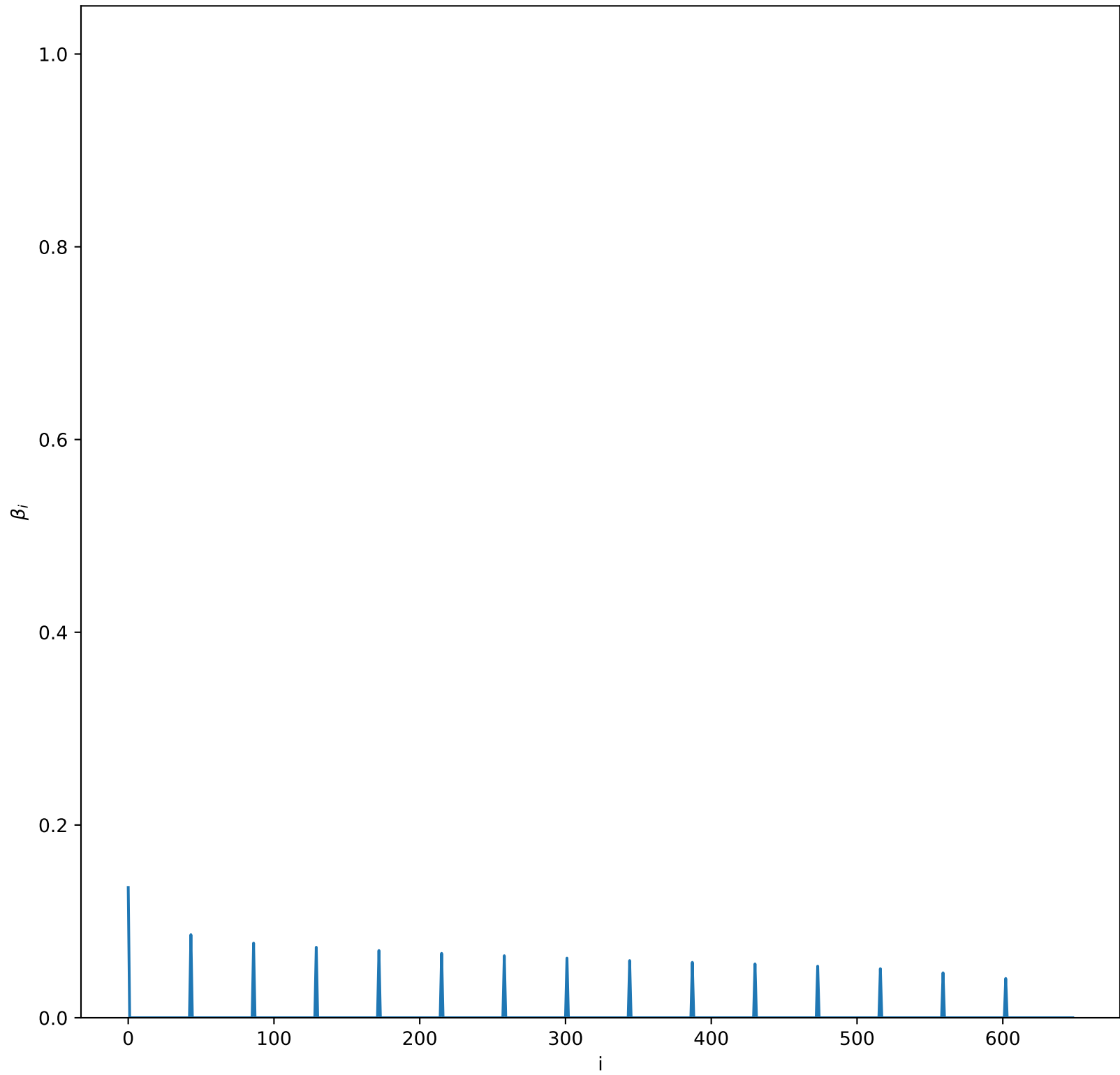
$\mu = 0.13$



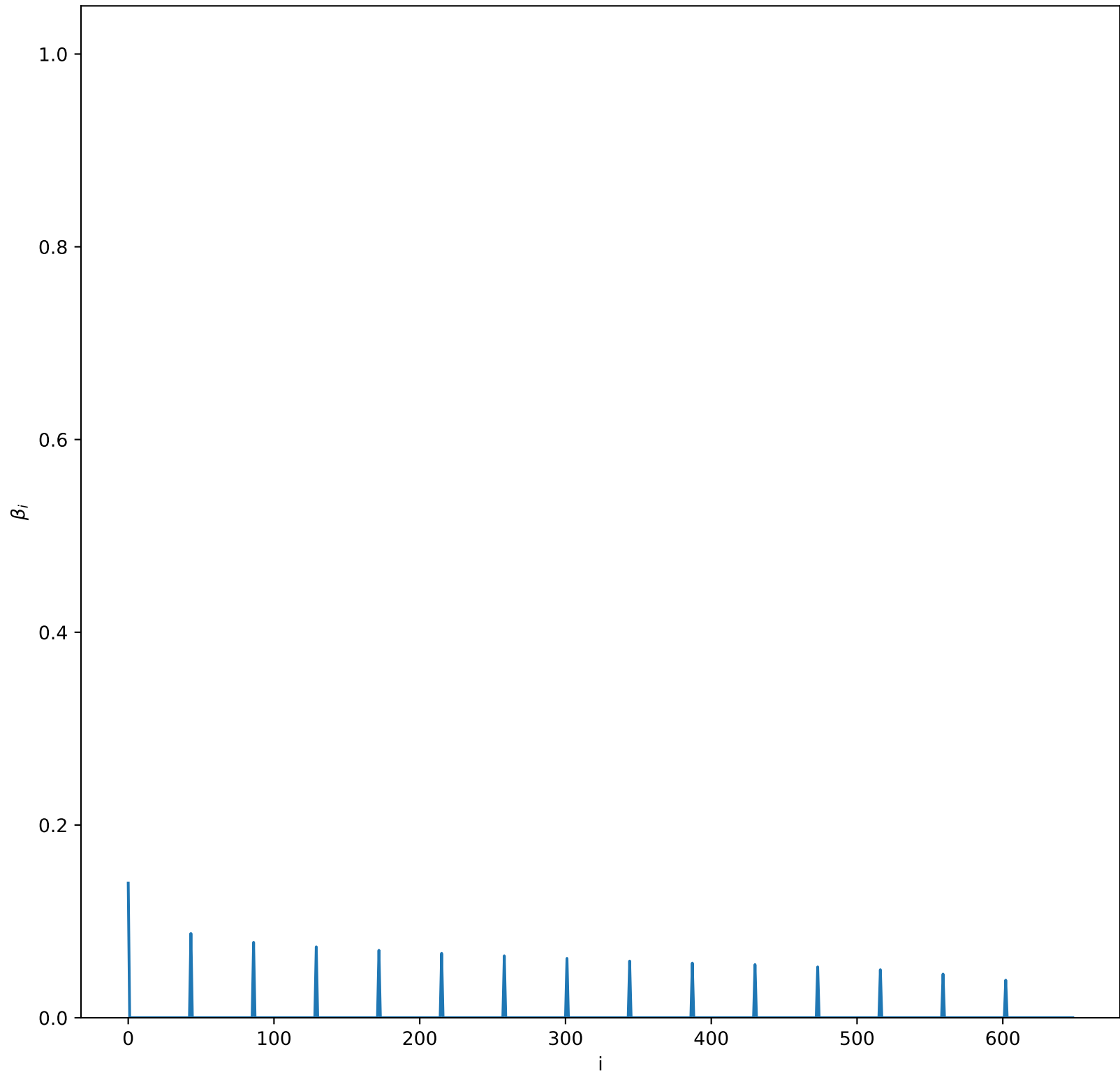
$\mu = 0.14$



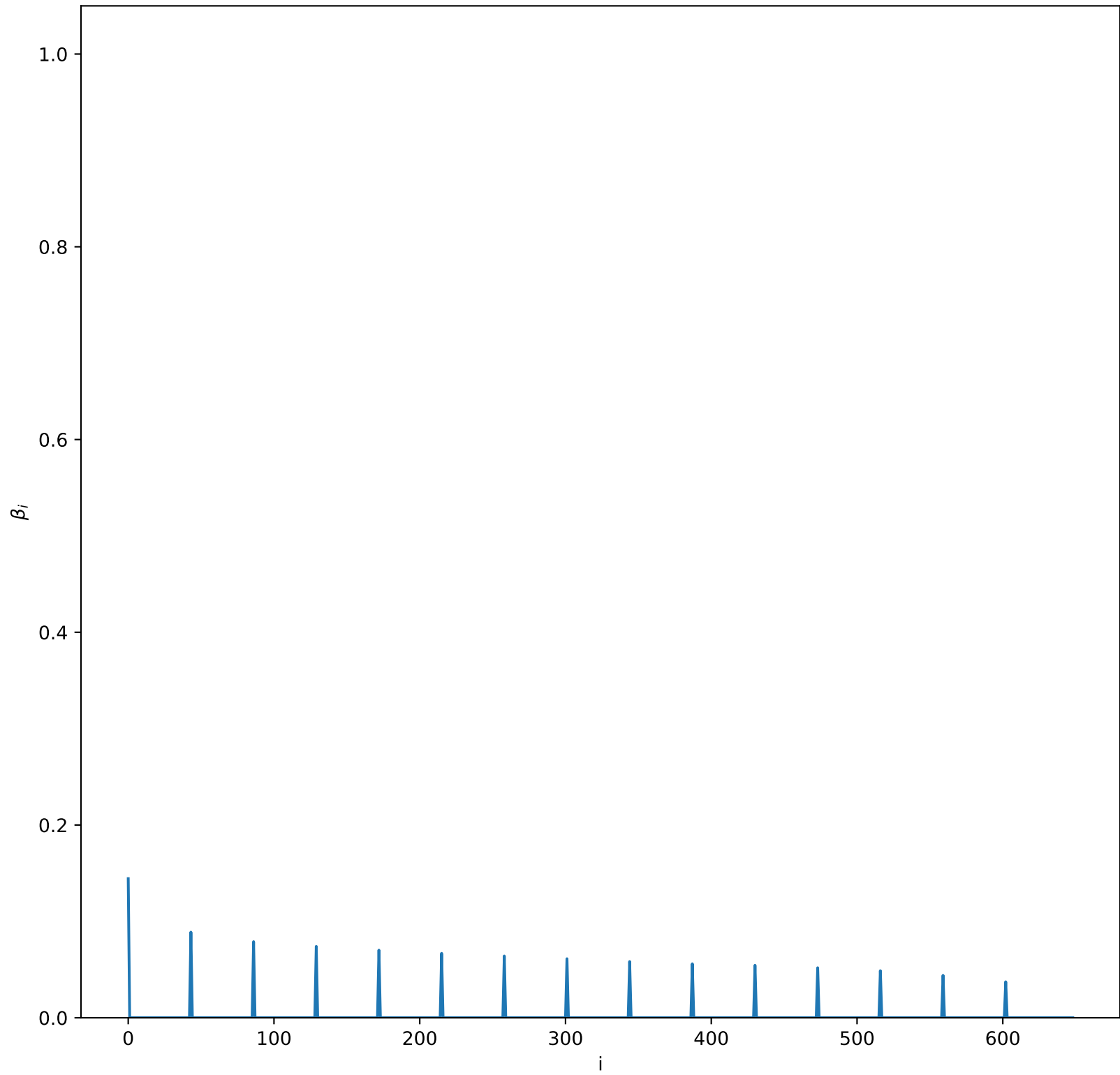
$\mu = 0.15$



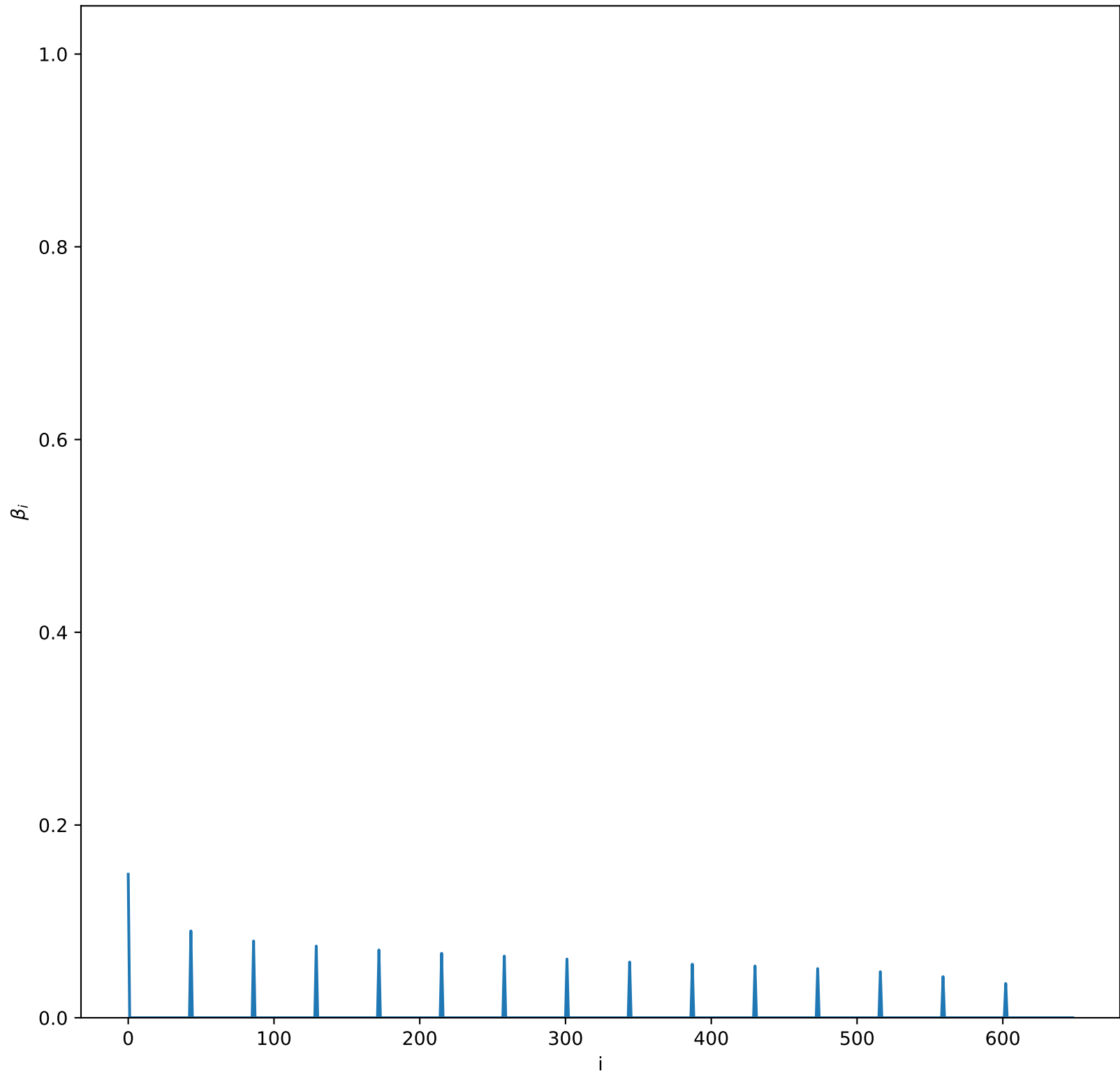
$\mu = 0.16$



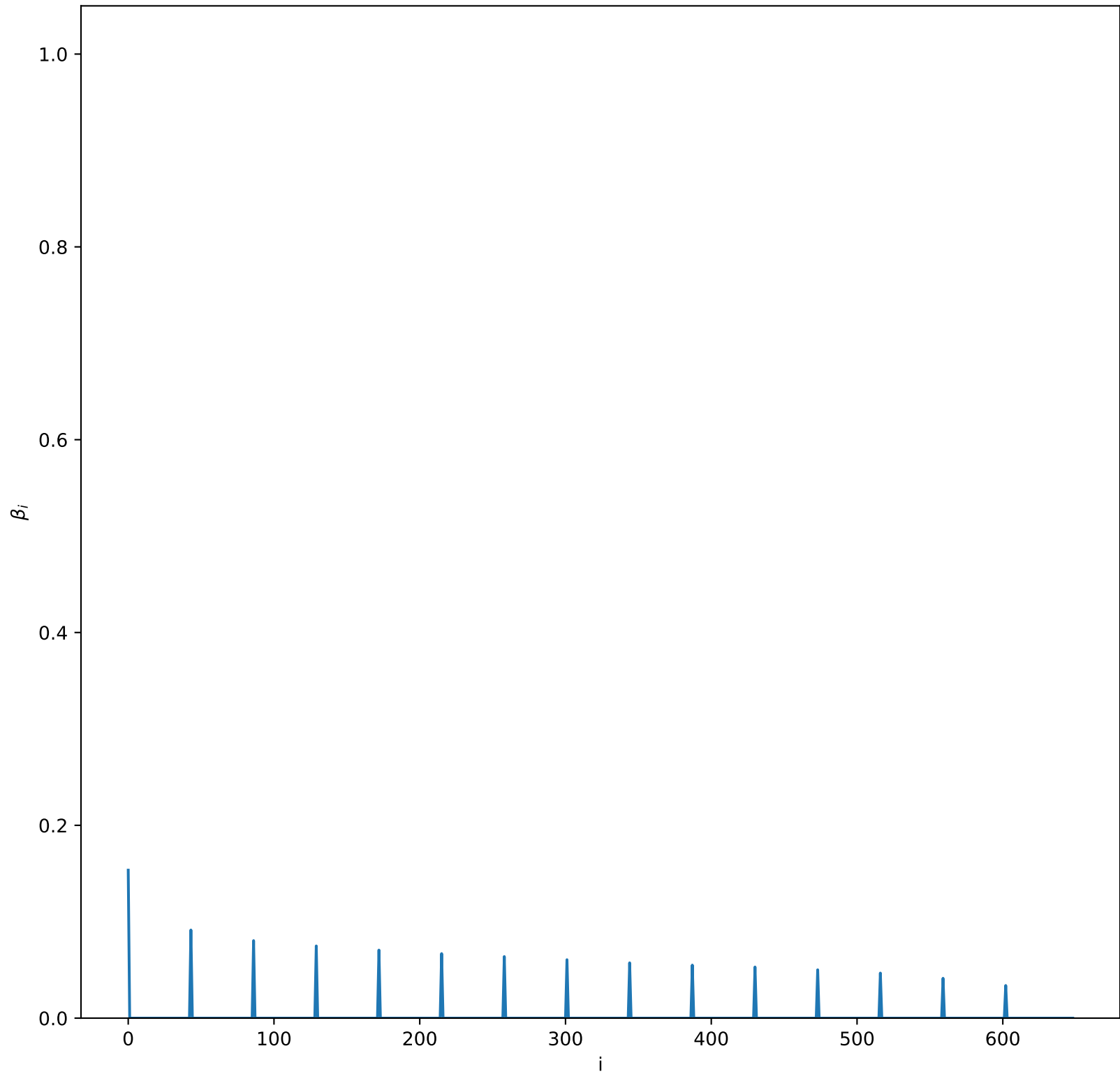
$\mu = 0.17$



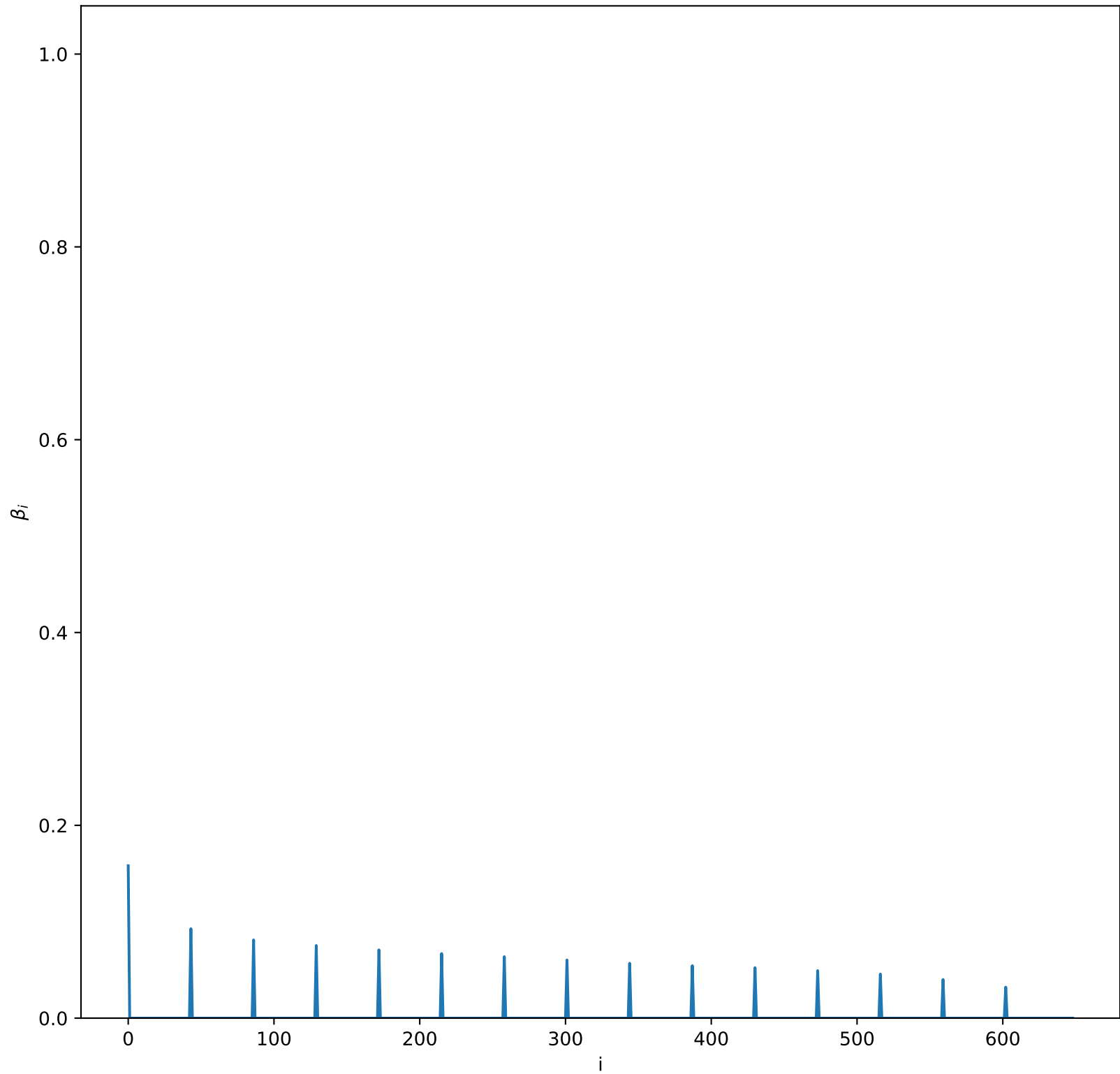
$\mu = 0.18$



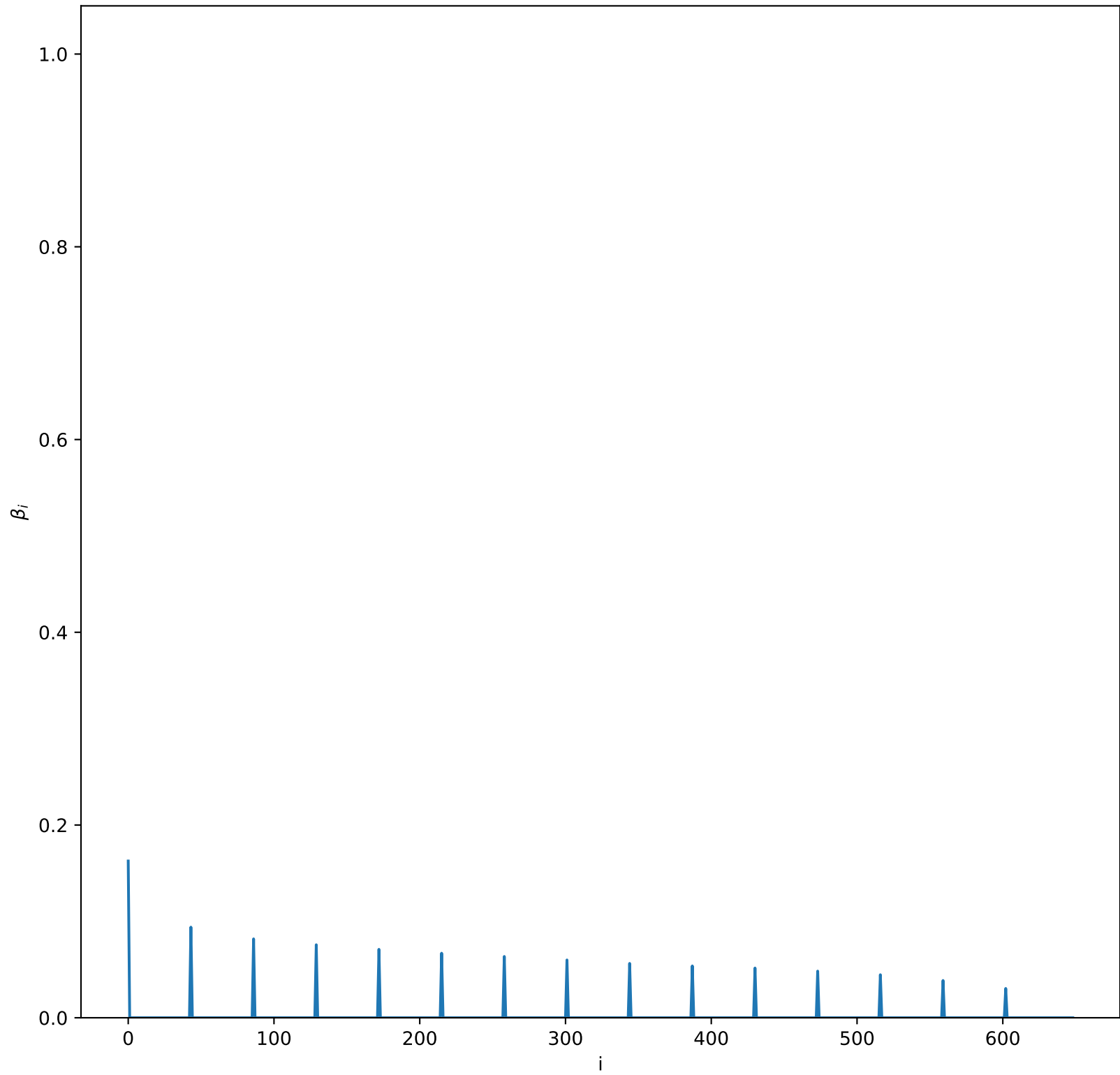
$\mu = 0.19$



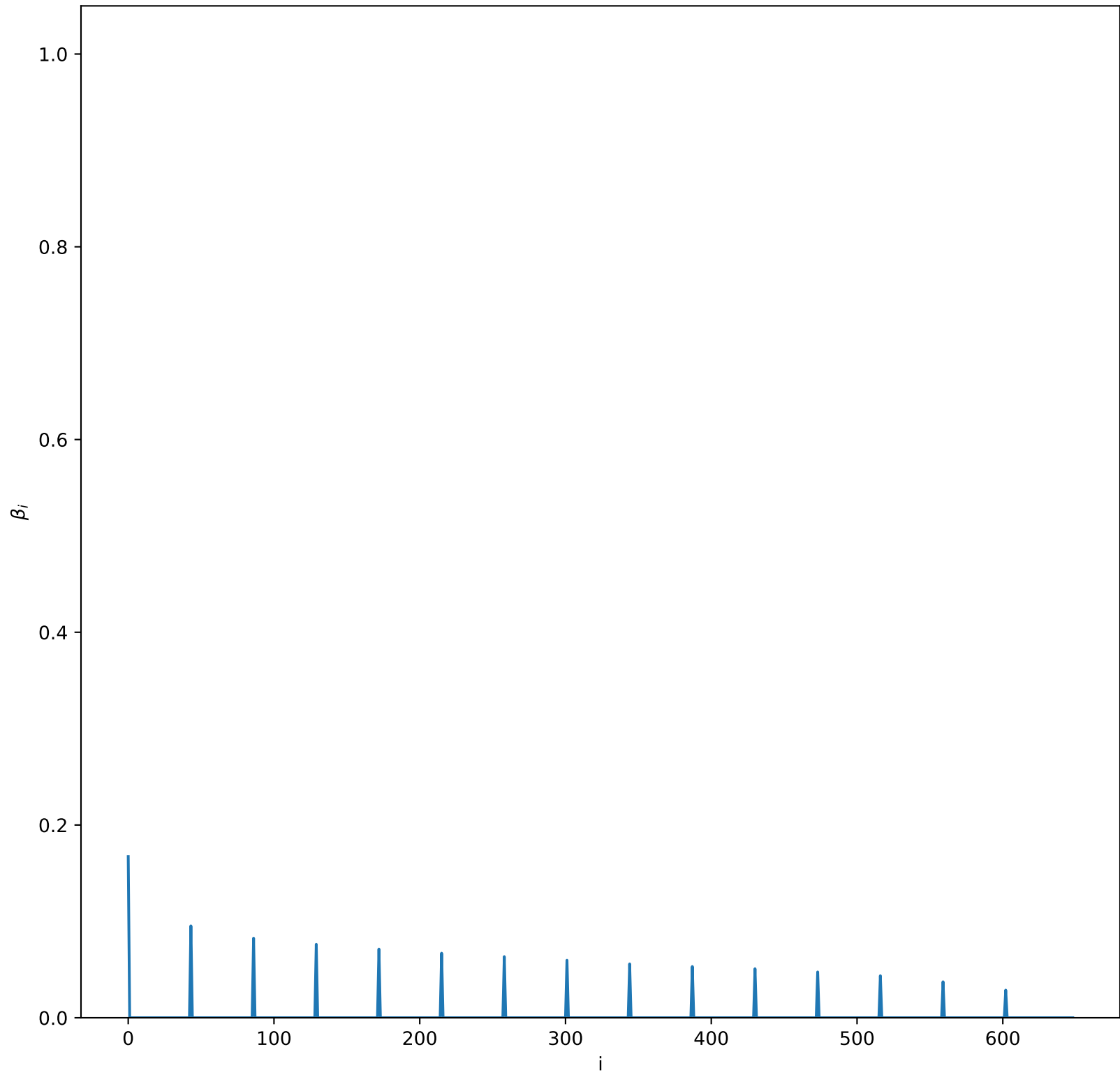
$\mu = 0.20$



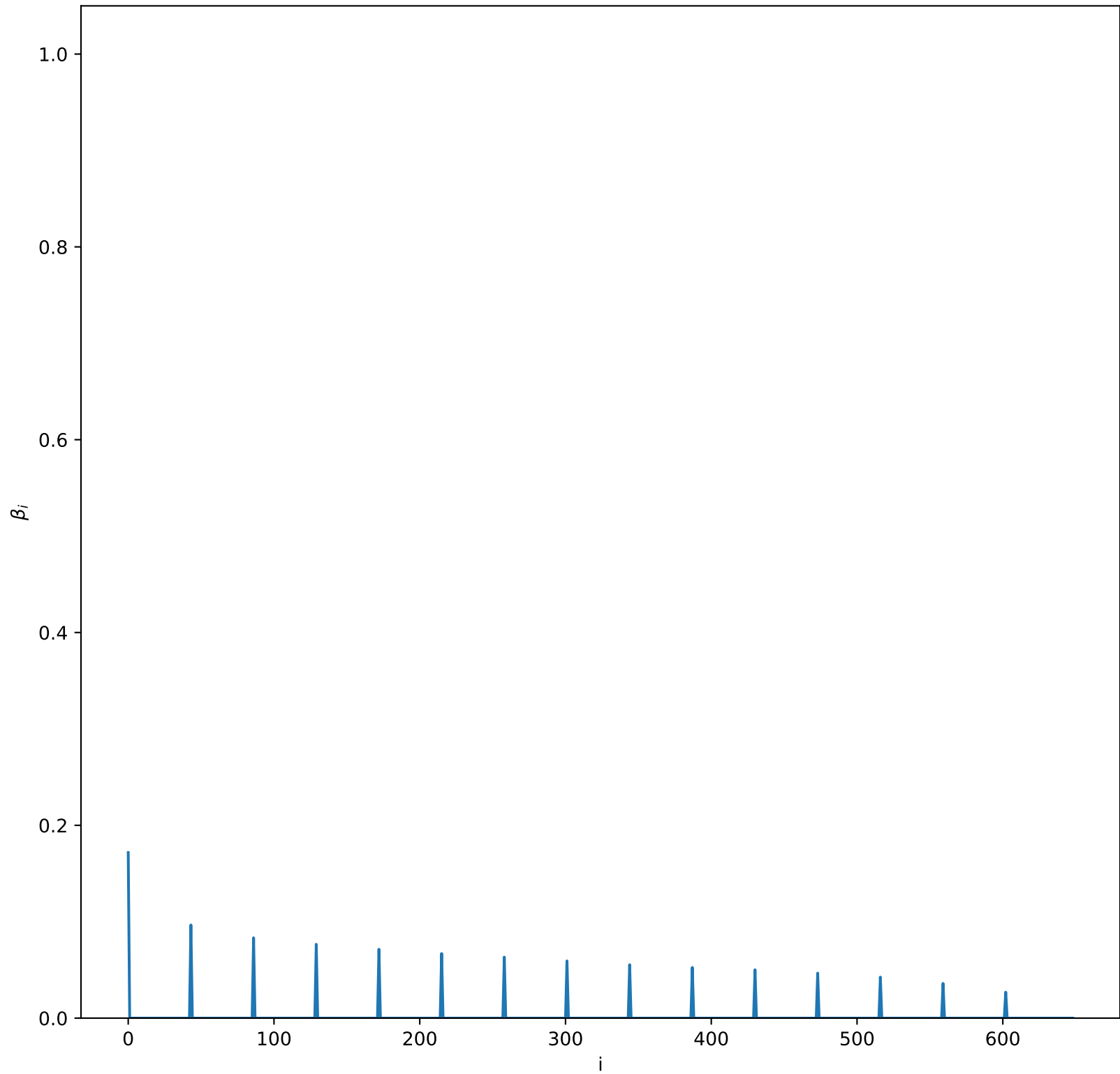
$\mu = 0.21$



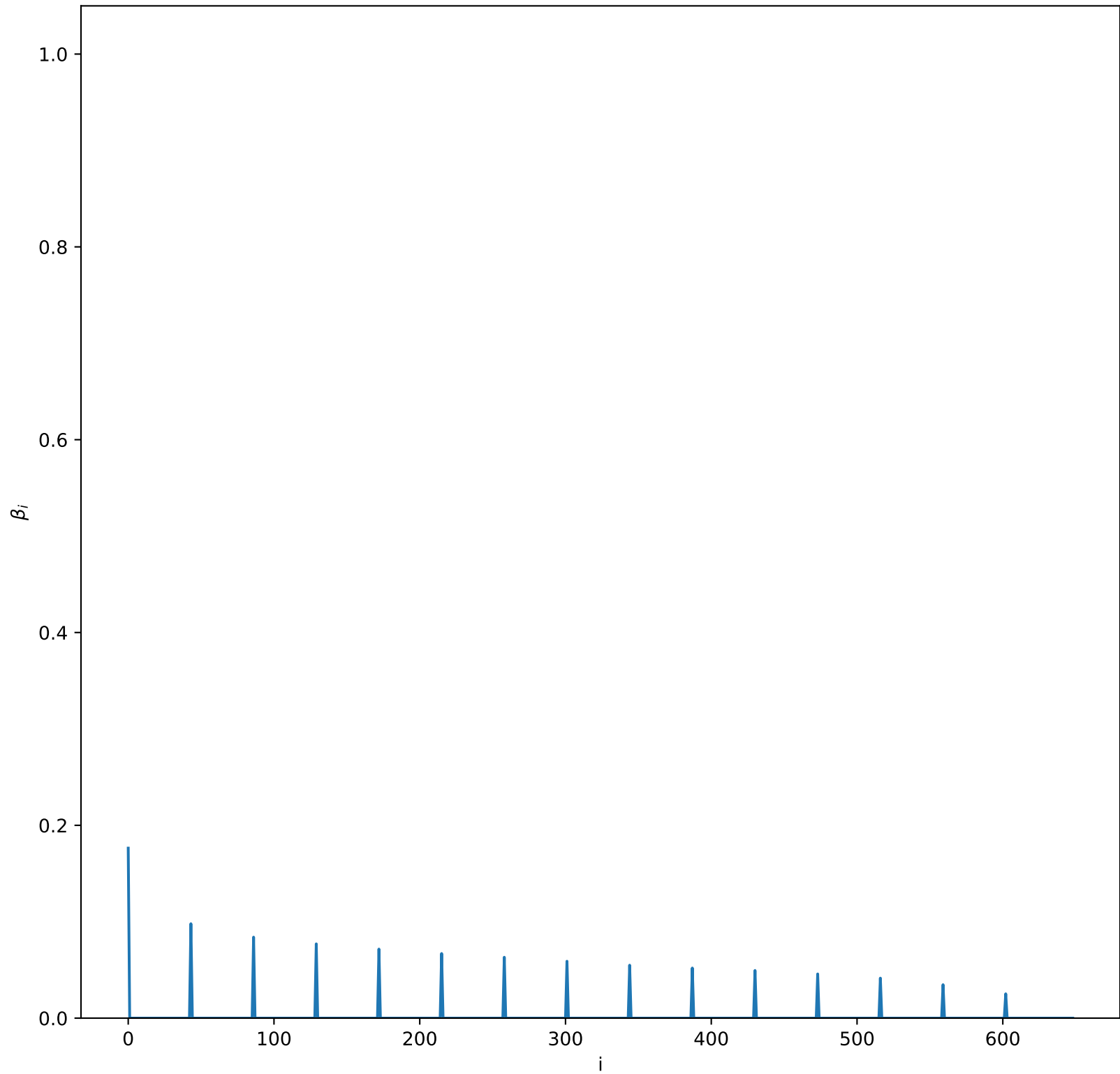
$\mu = 0.22$



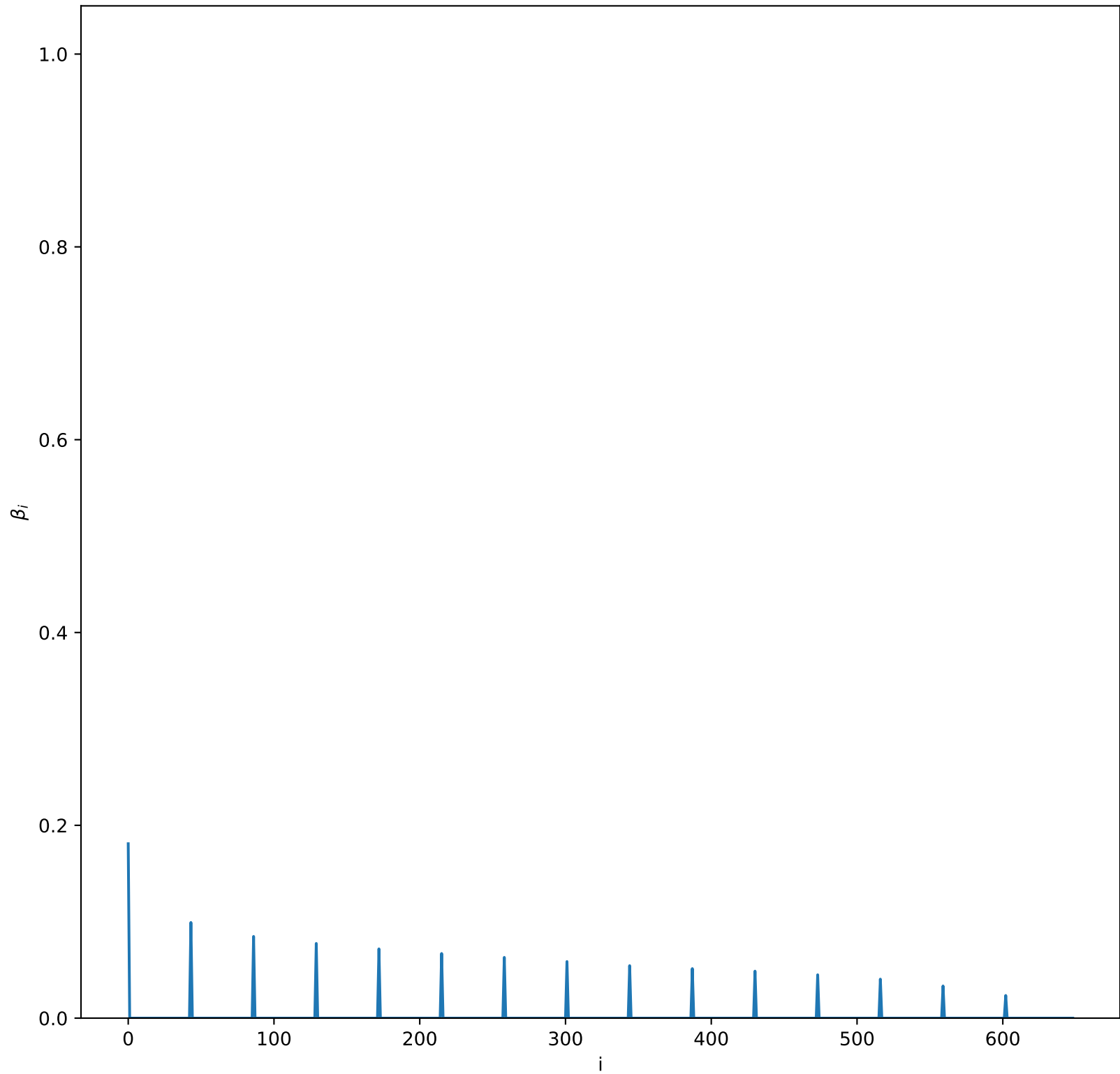
$\mu = 0.23$



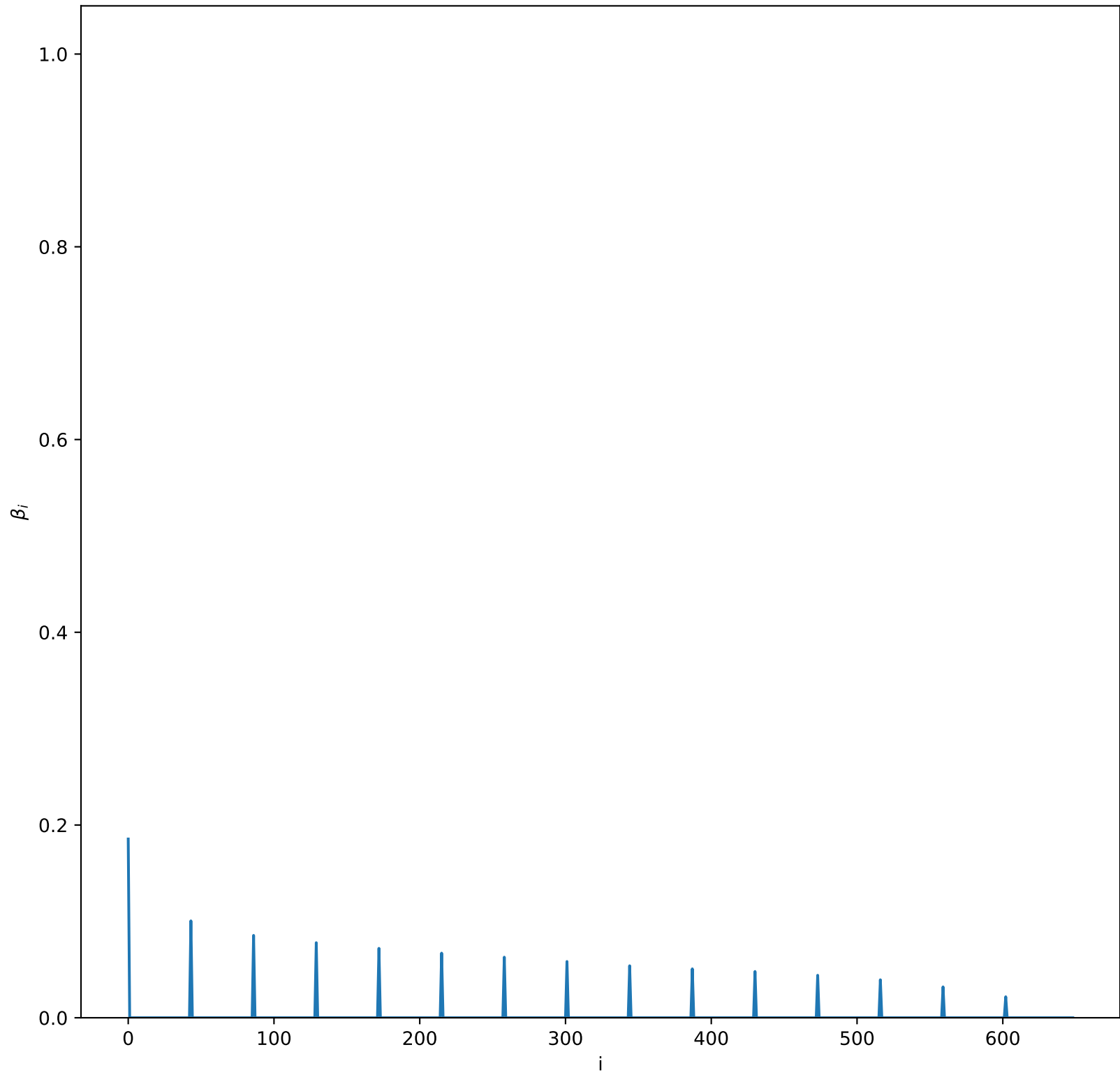
$\mu = 0.24$



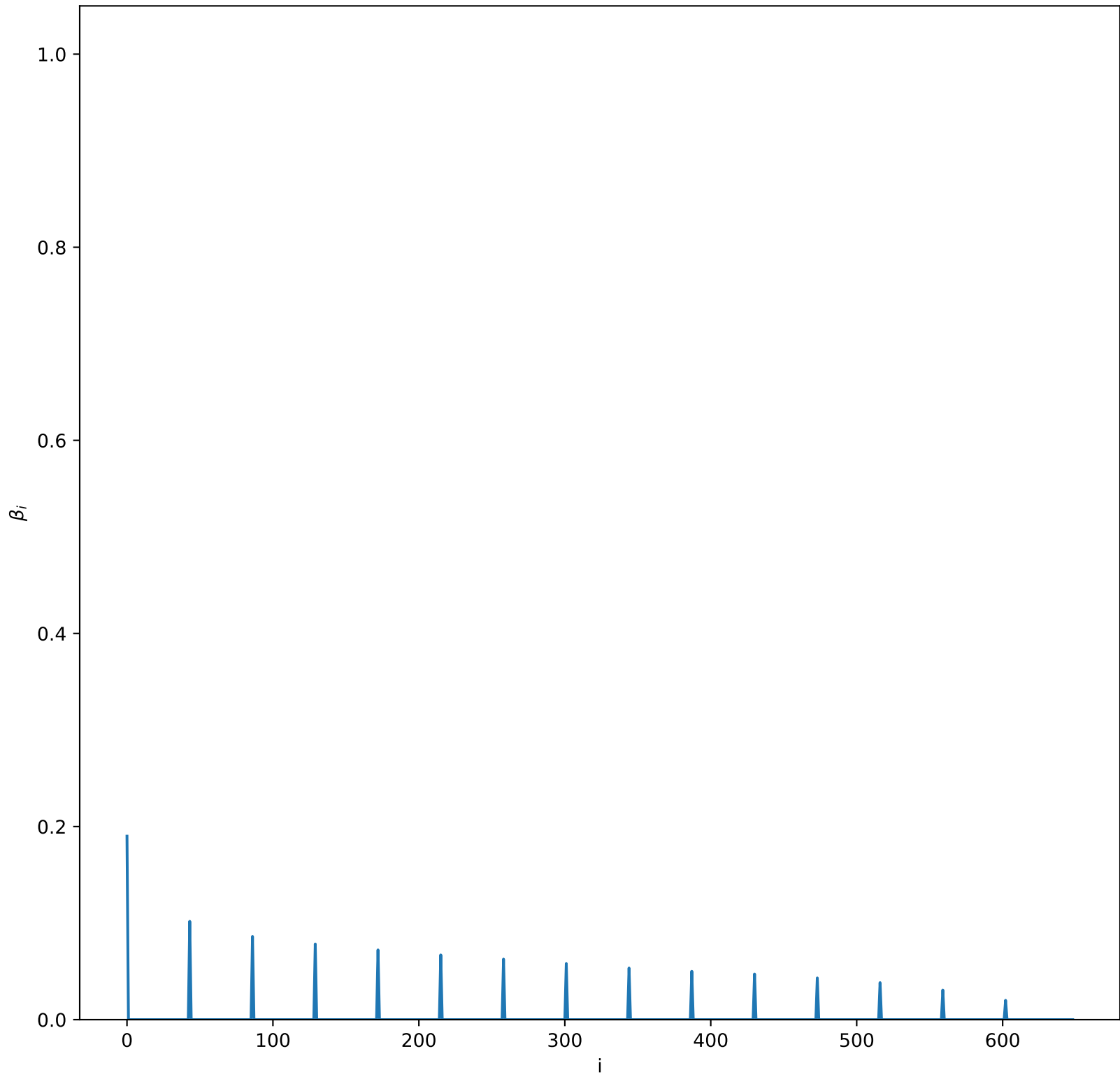
$\mu = 0.25$



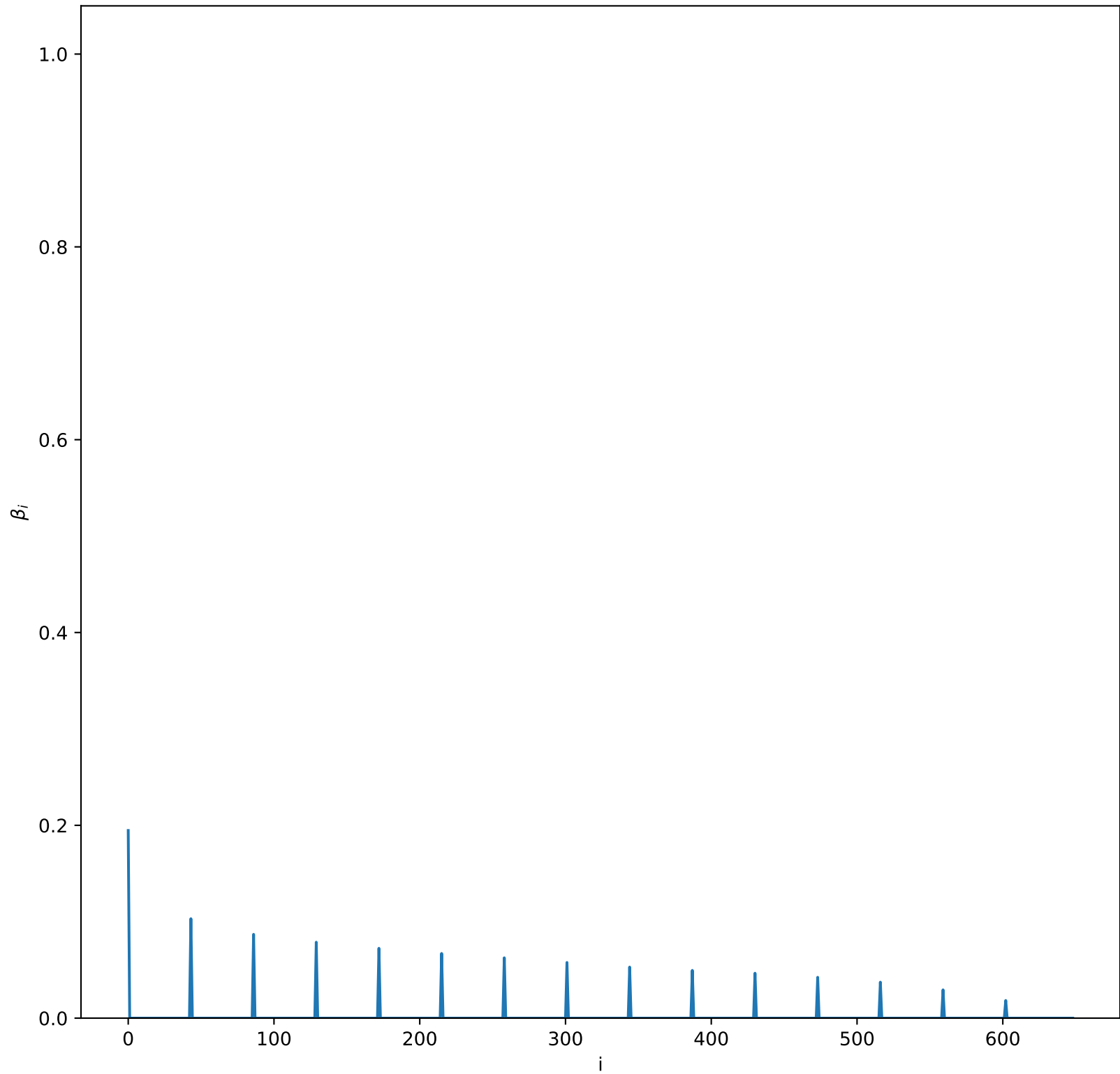
$\mu = 0.26$



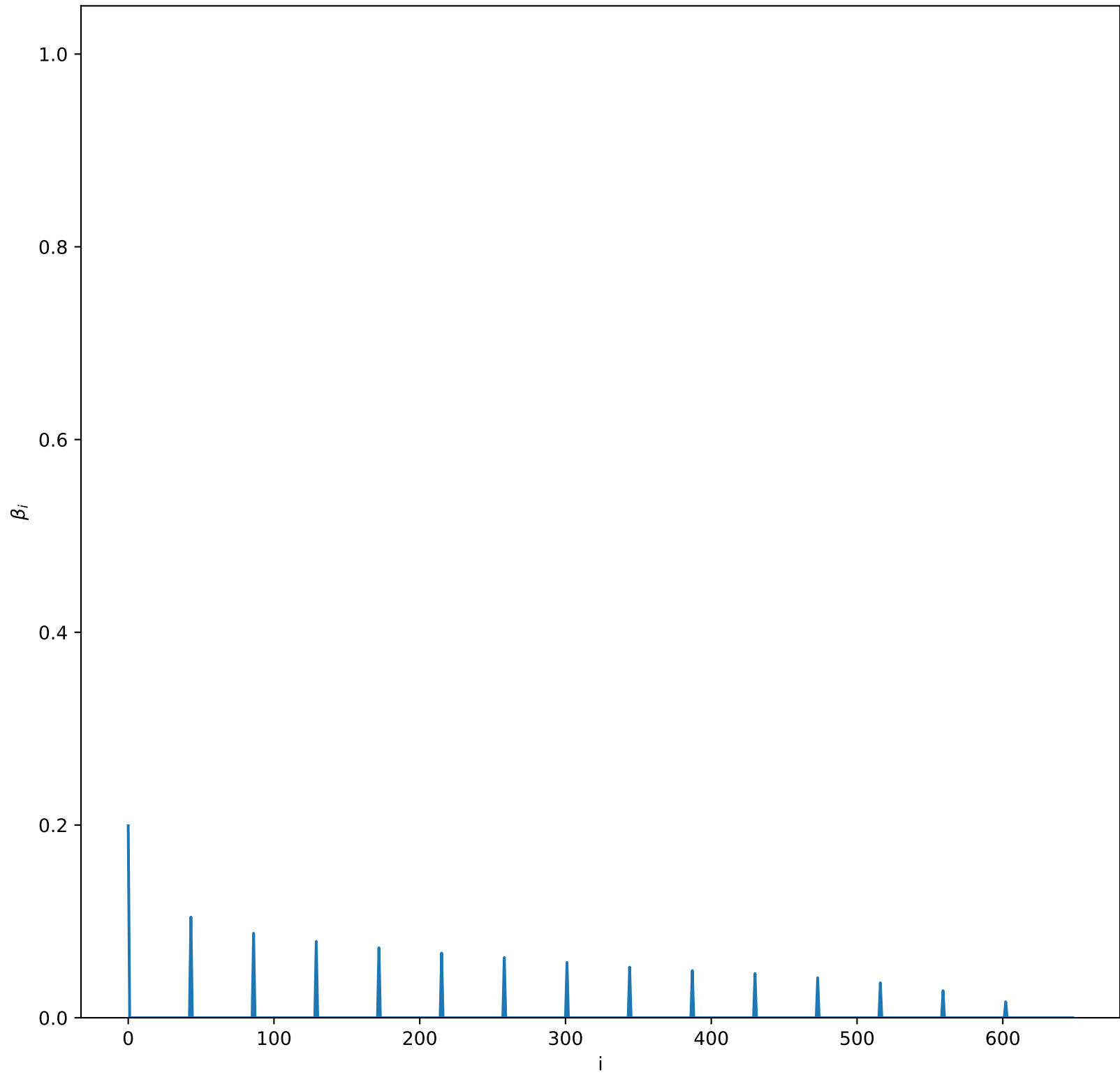
$\mu = 0.27$



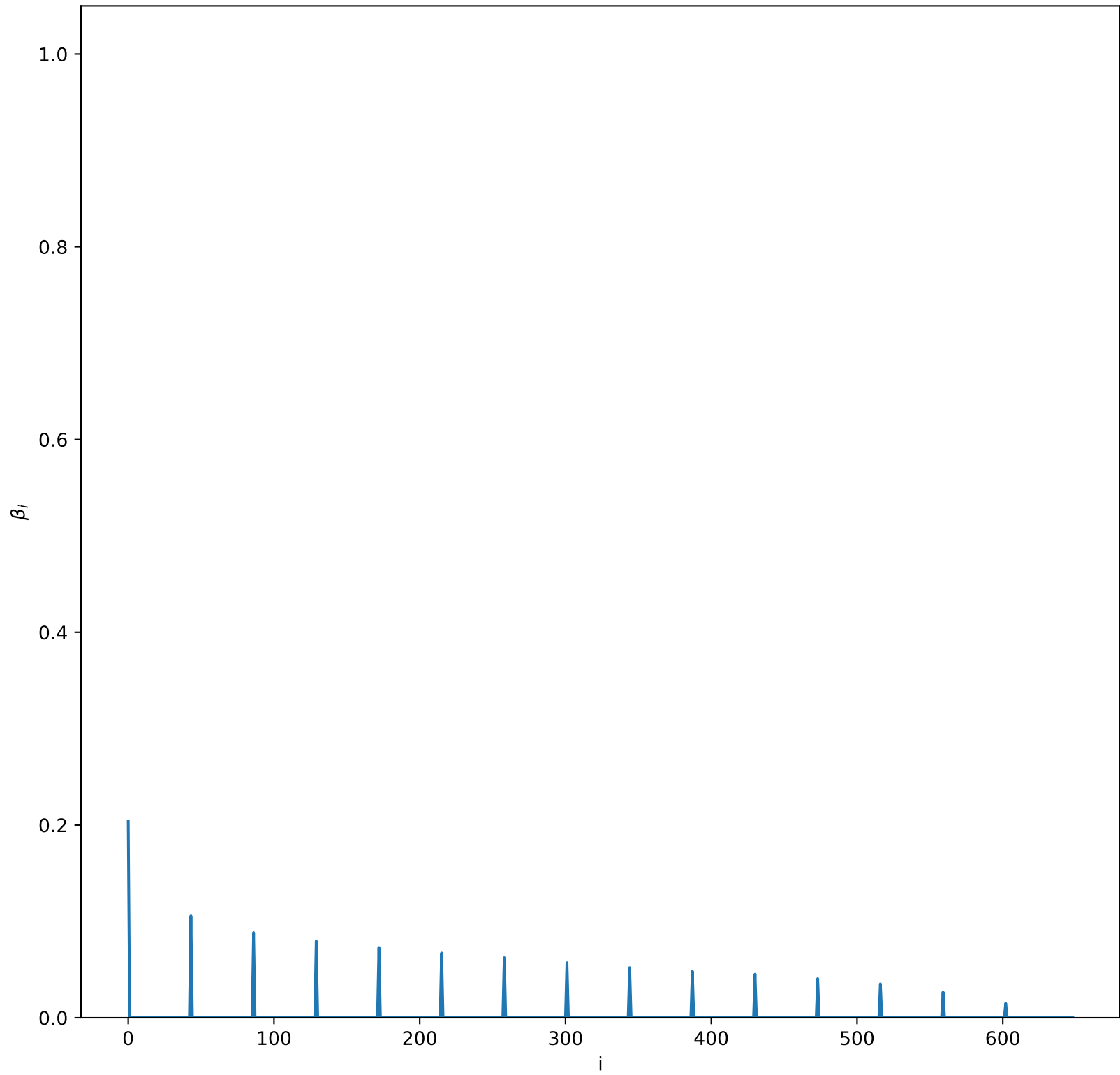
$\mu = 0.28$



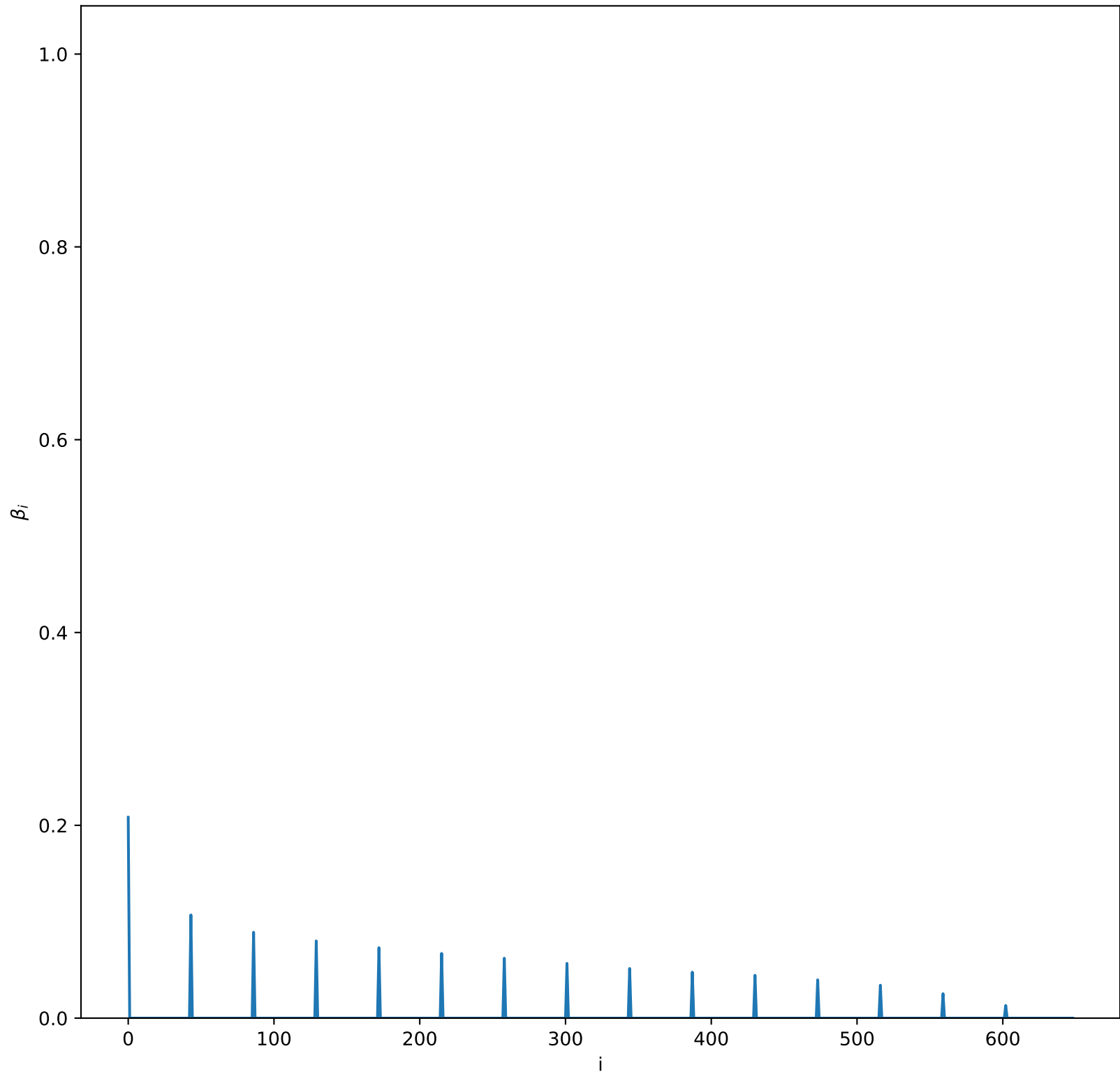
$\mu = 0.29$



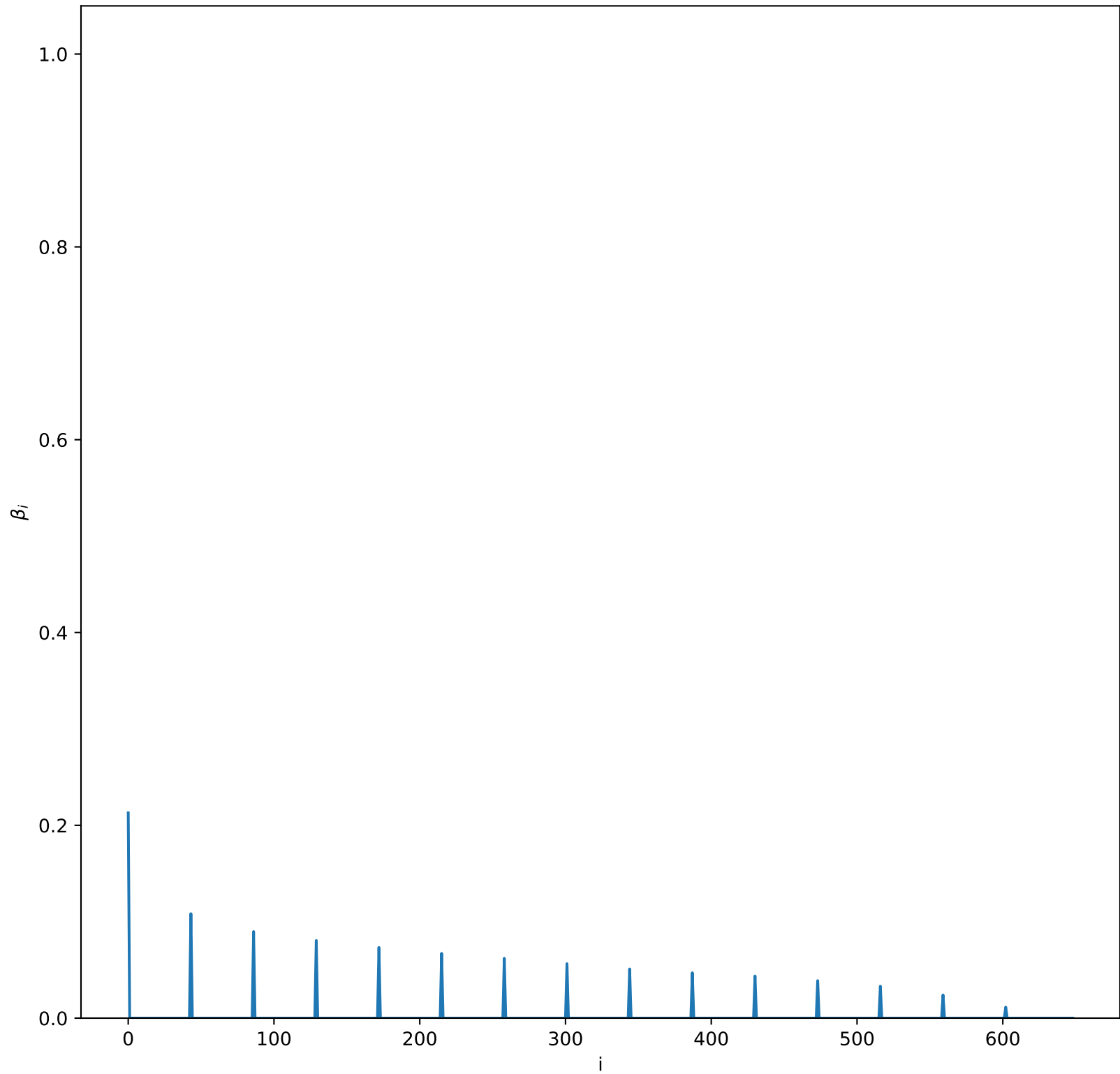
$\mu = 0.30$



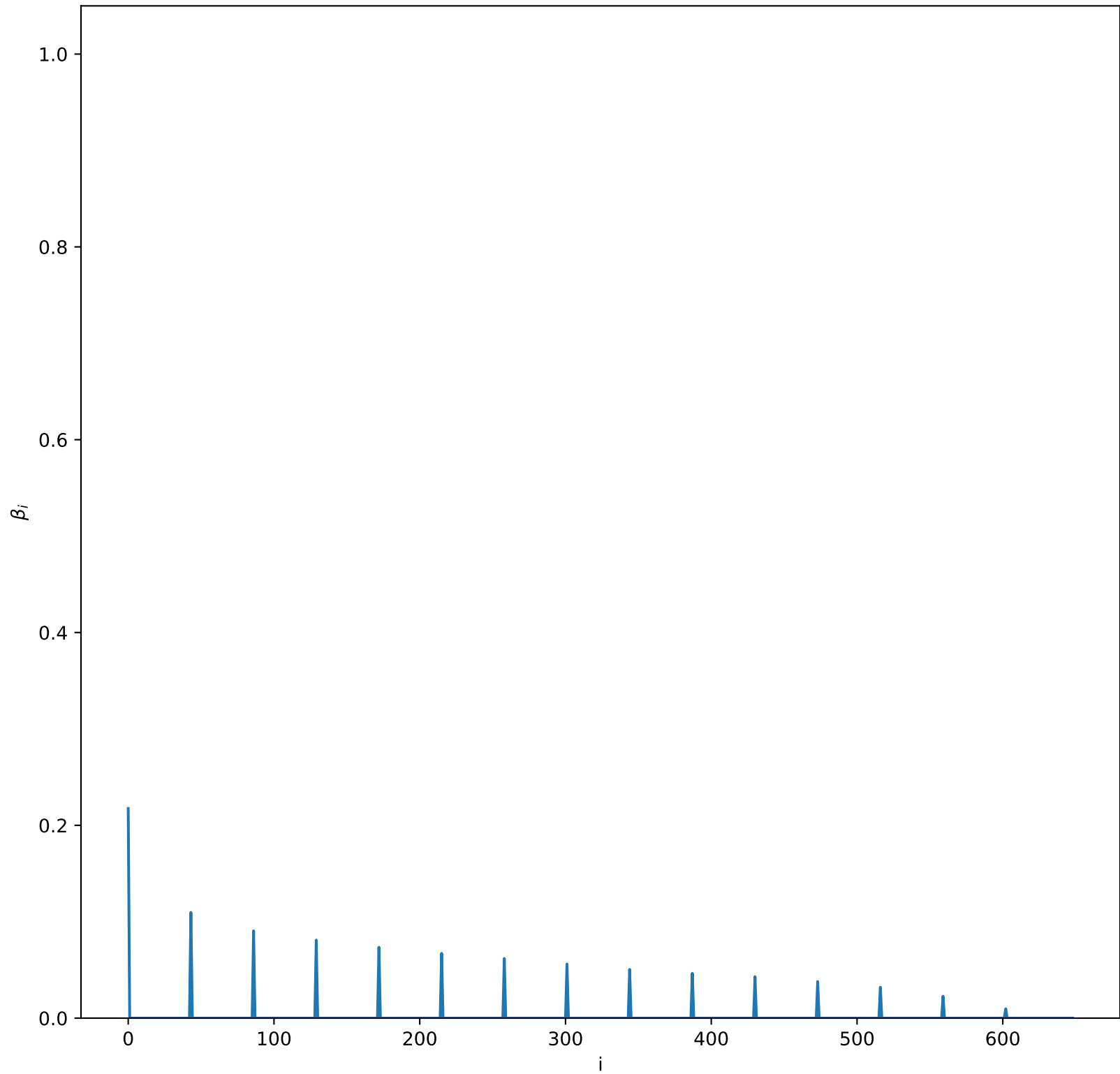
$\mu = 0.31$



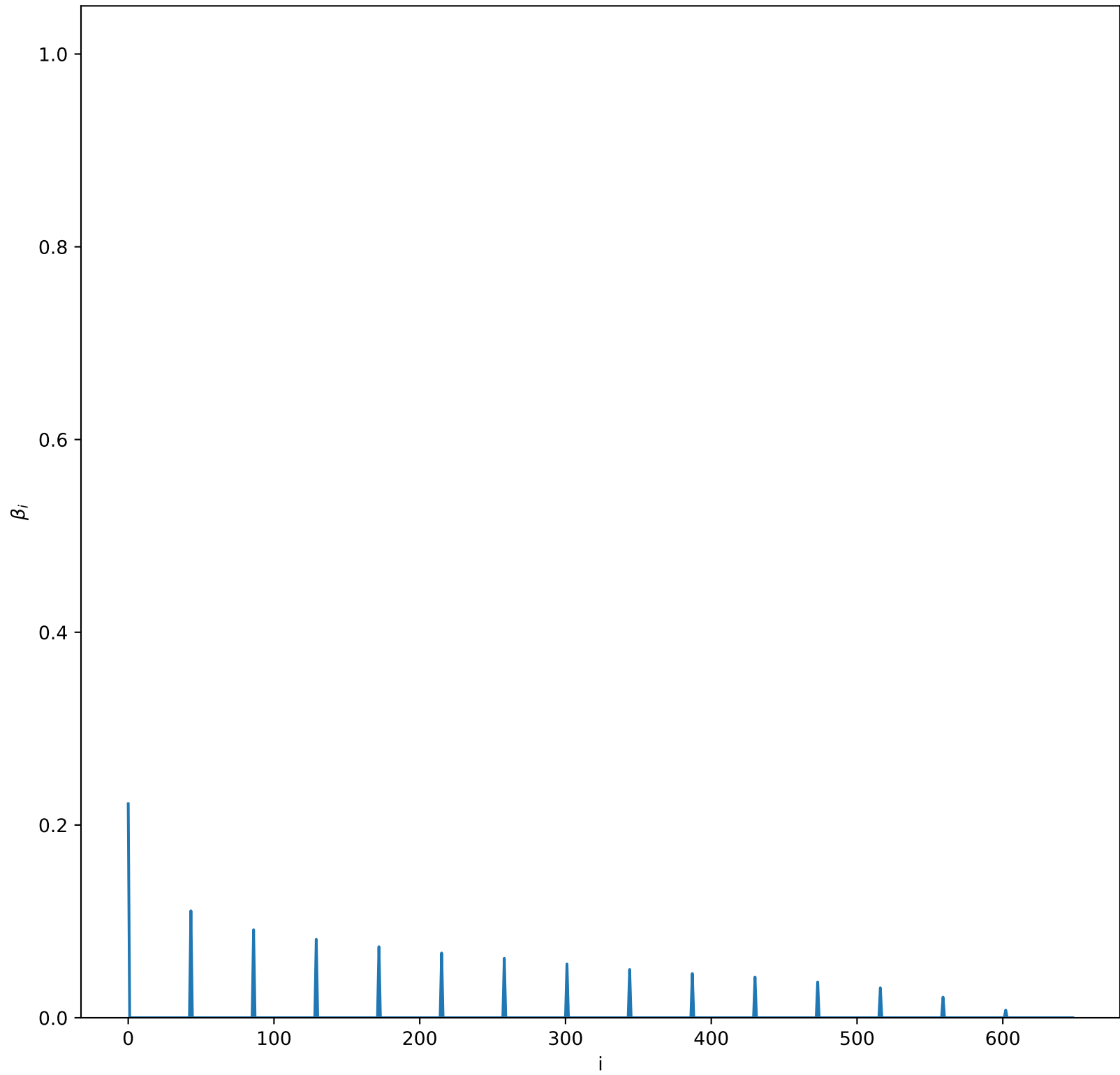
$\mu = 0.32$



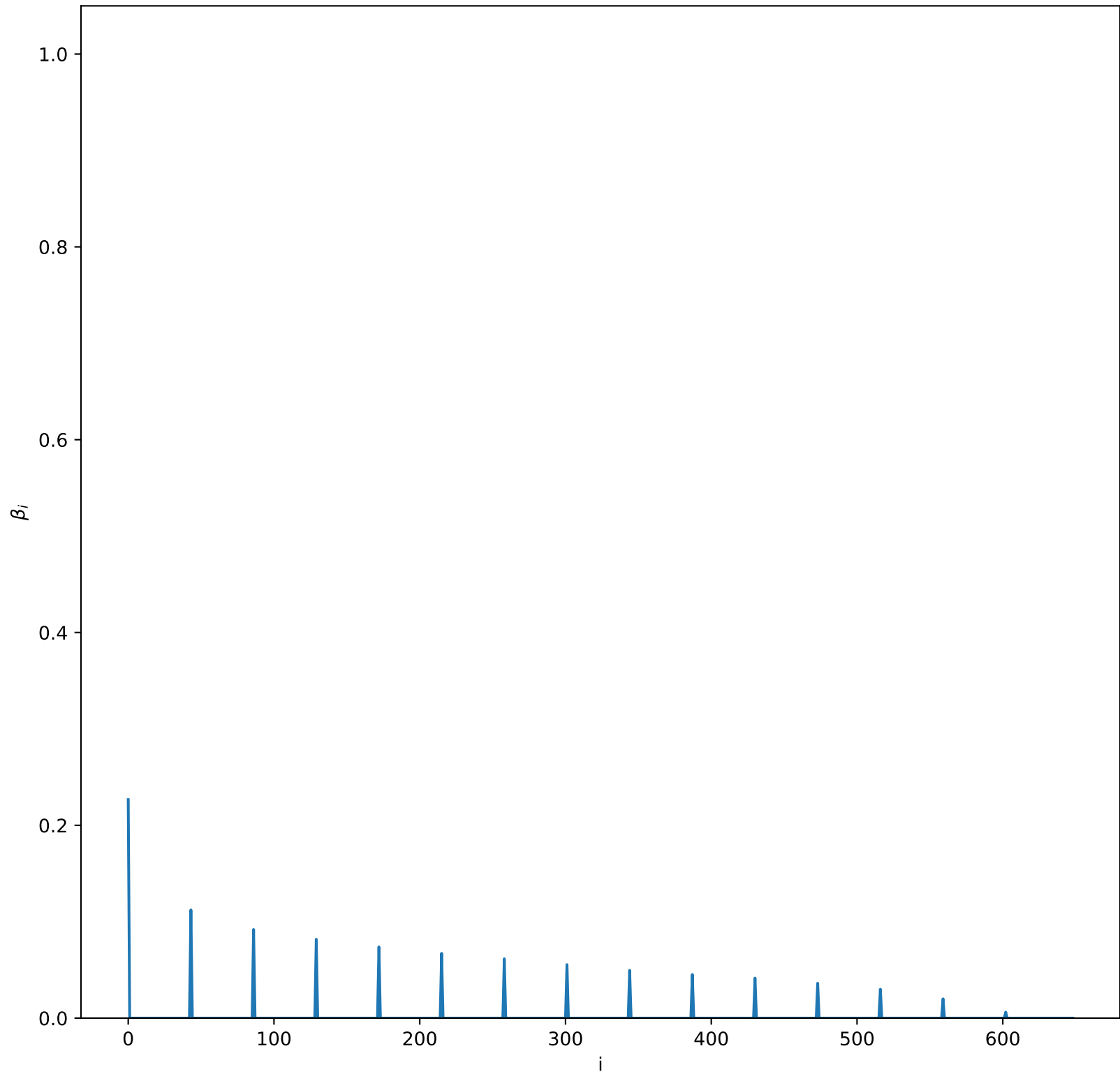
$\mu = 0.33$



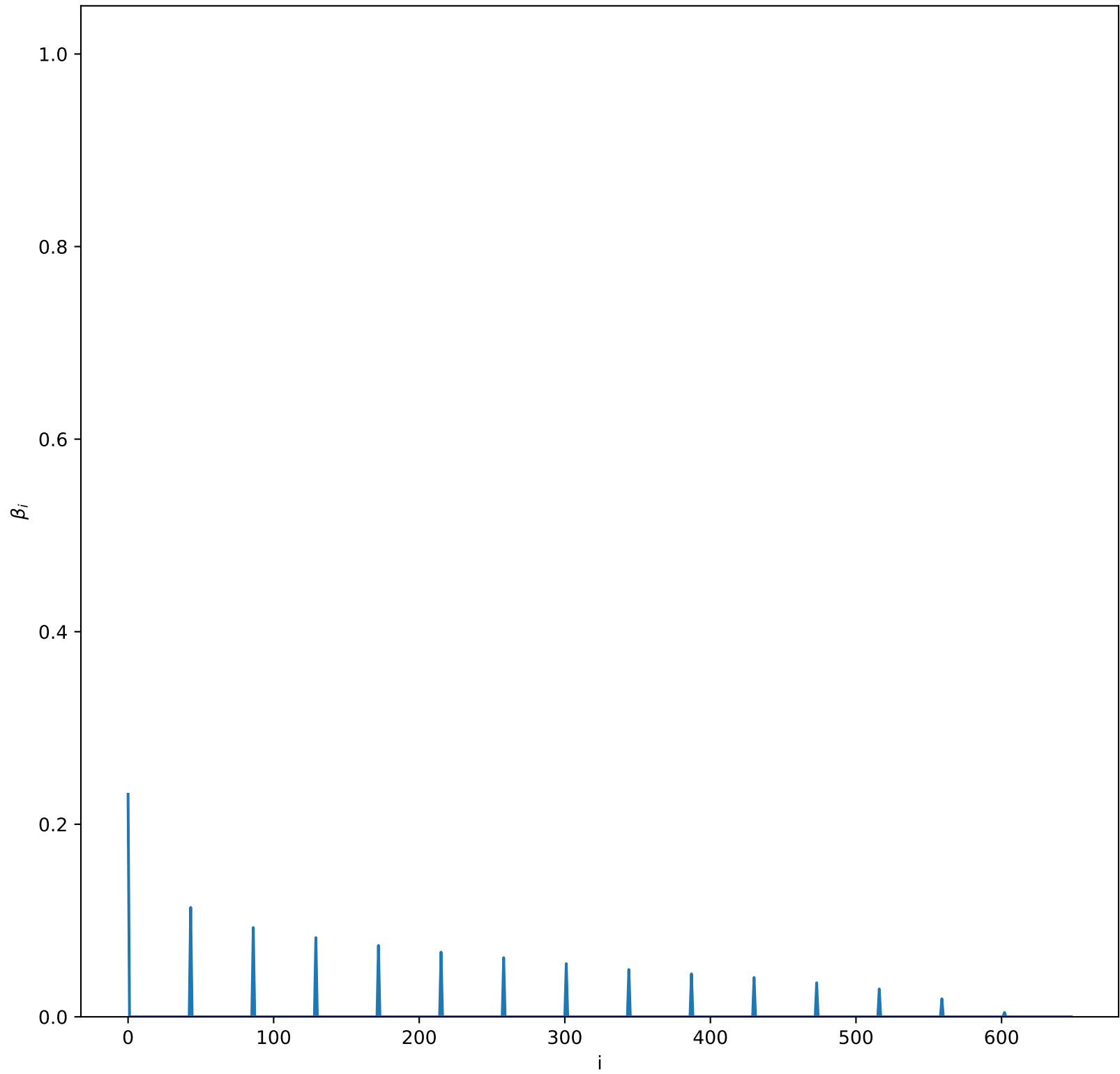
$\mu = 0.34$



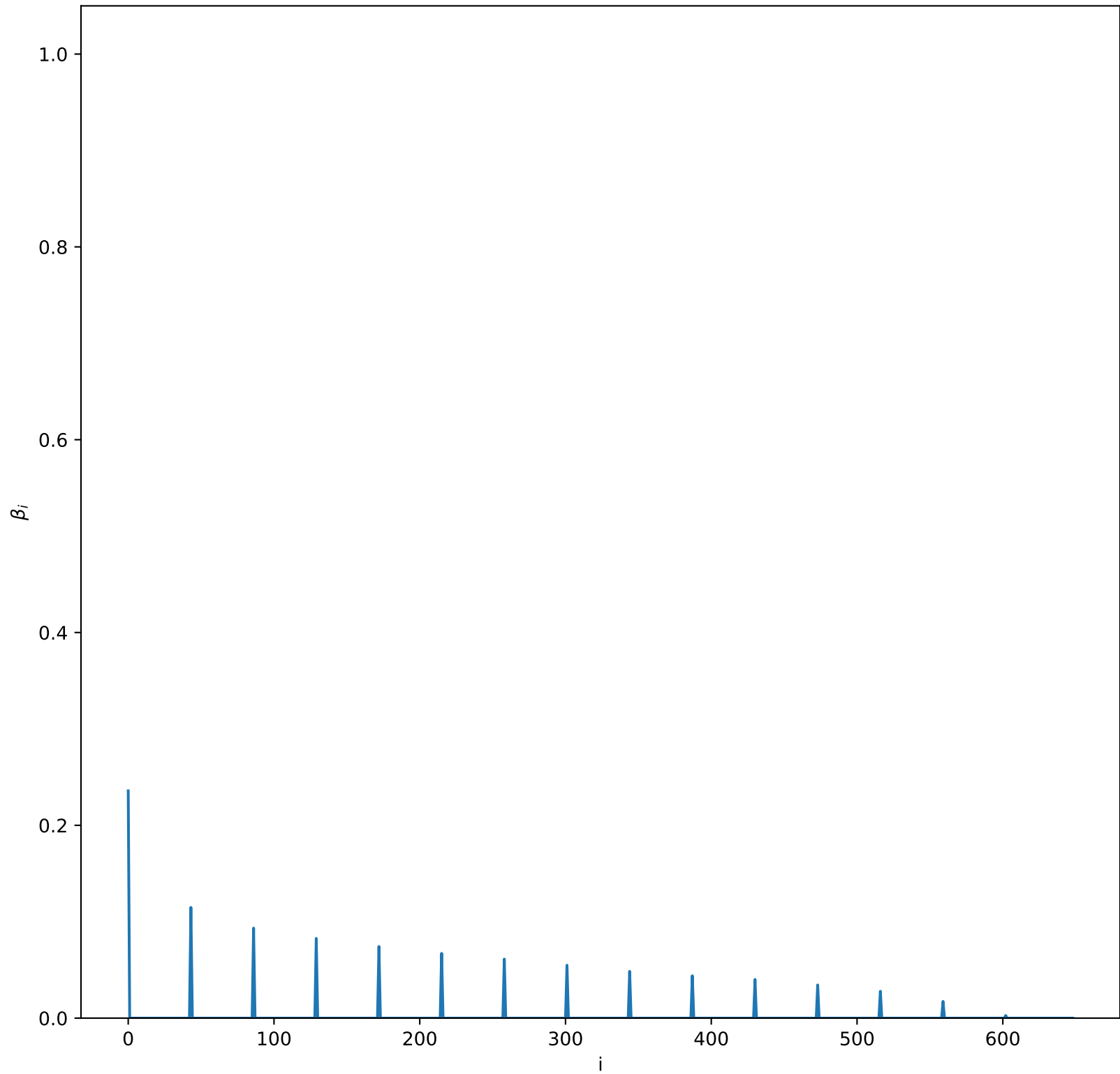
$\mu = 0.35$



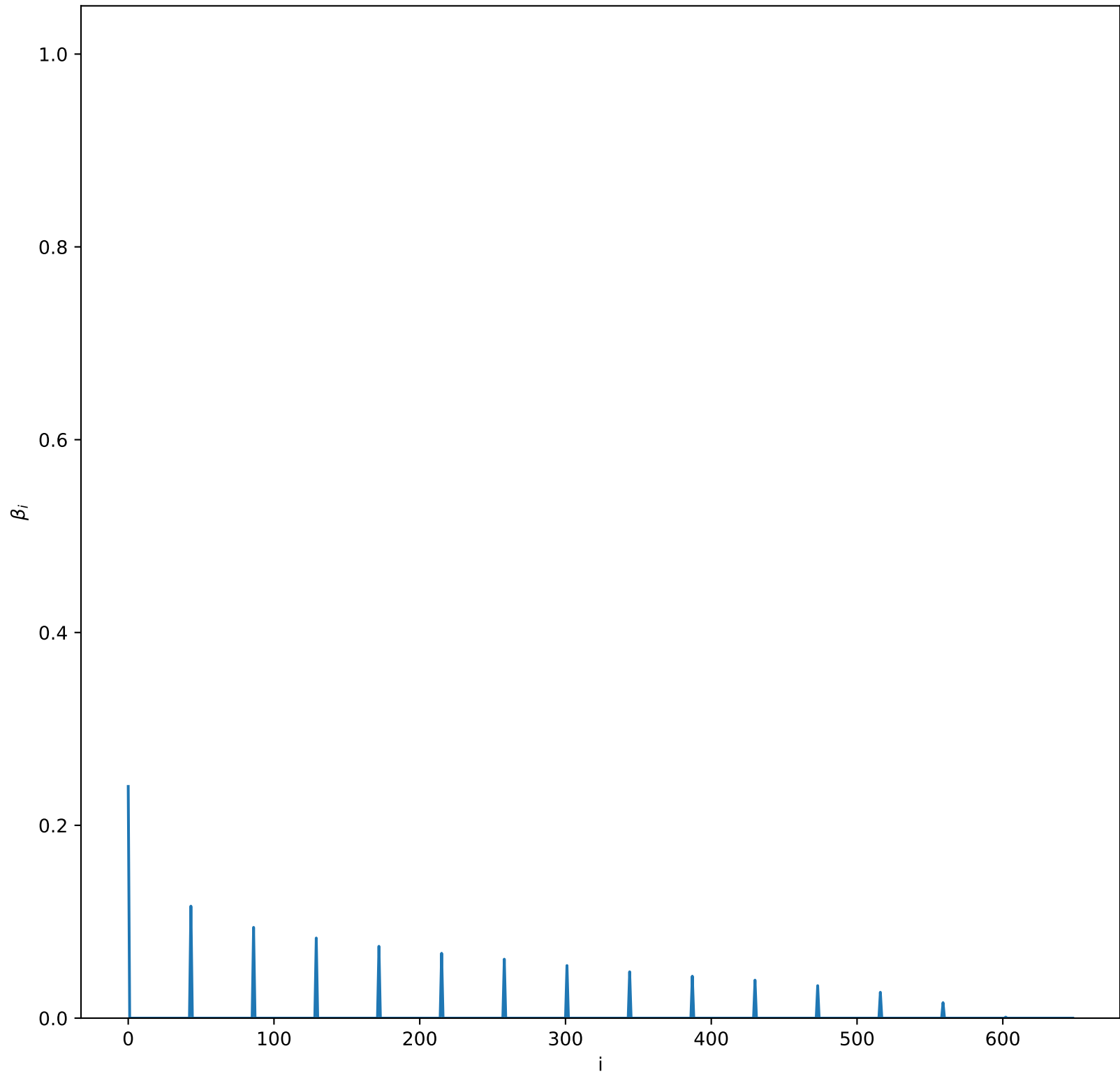
$\mu = 0.36$



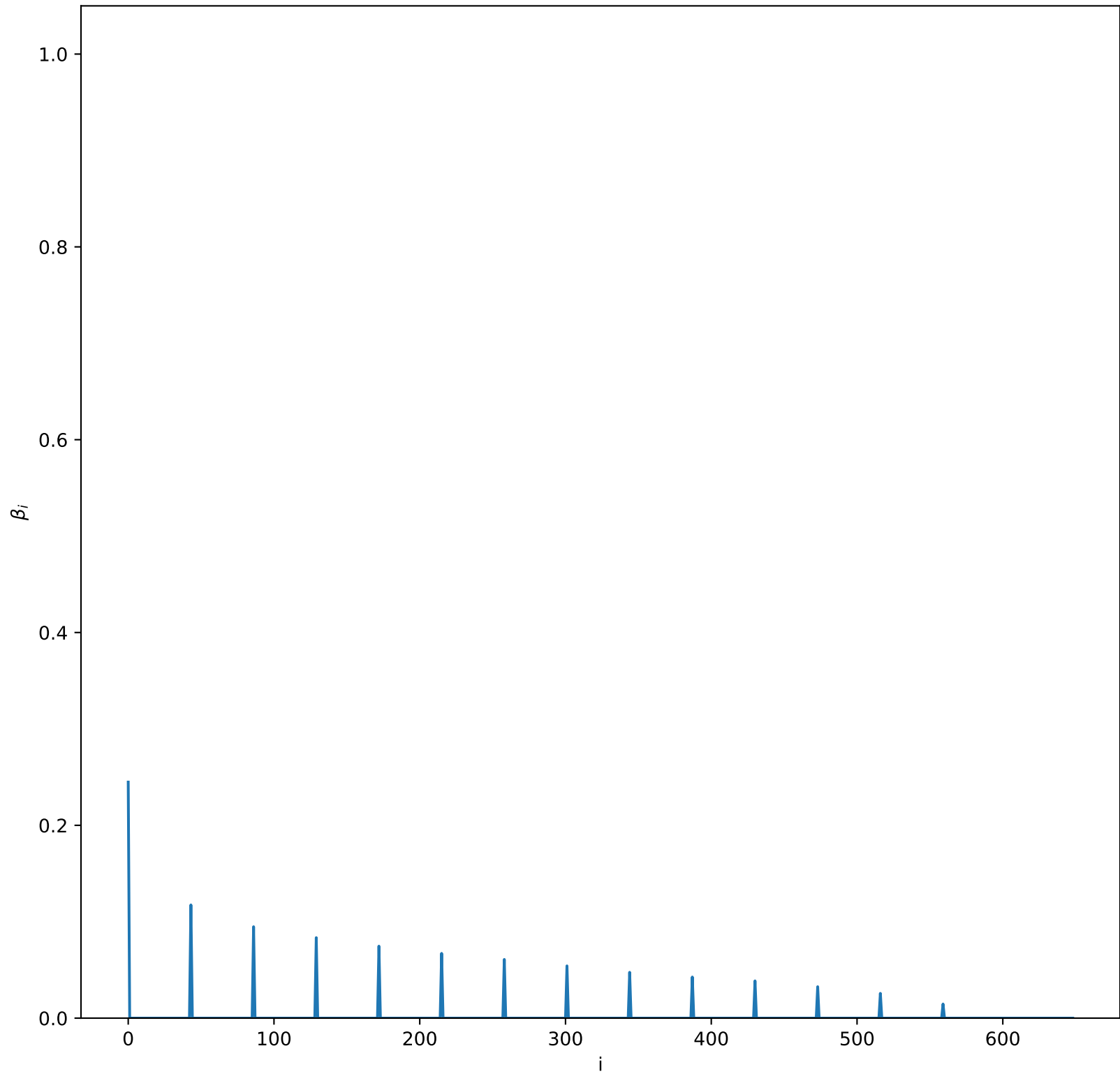
$\mu = 0.37$



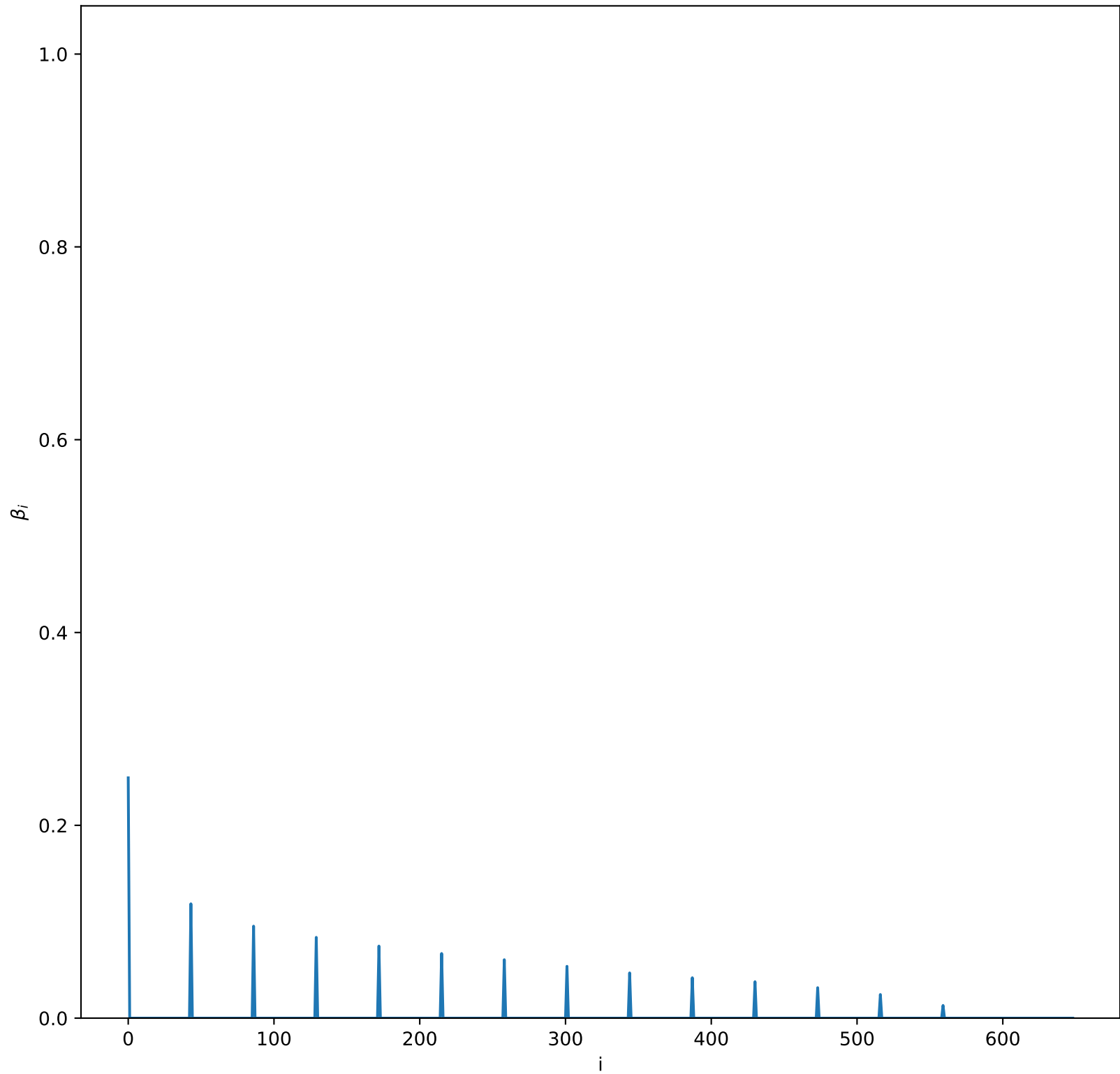
$\mu = 0.38$



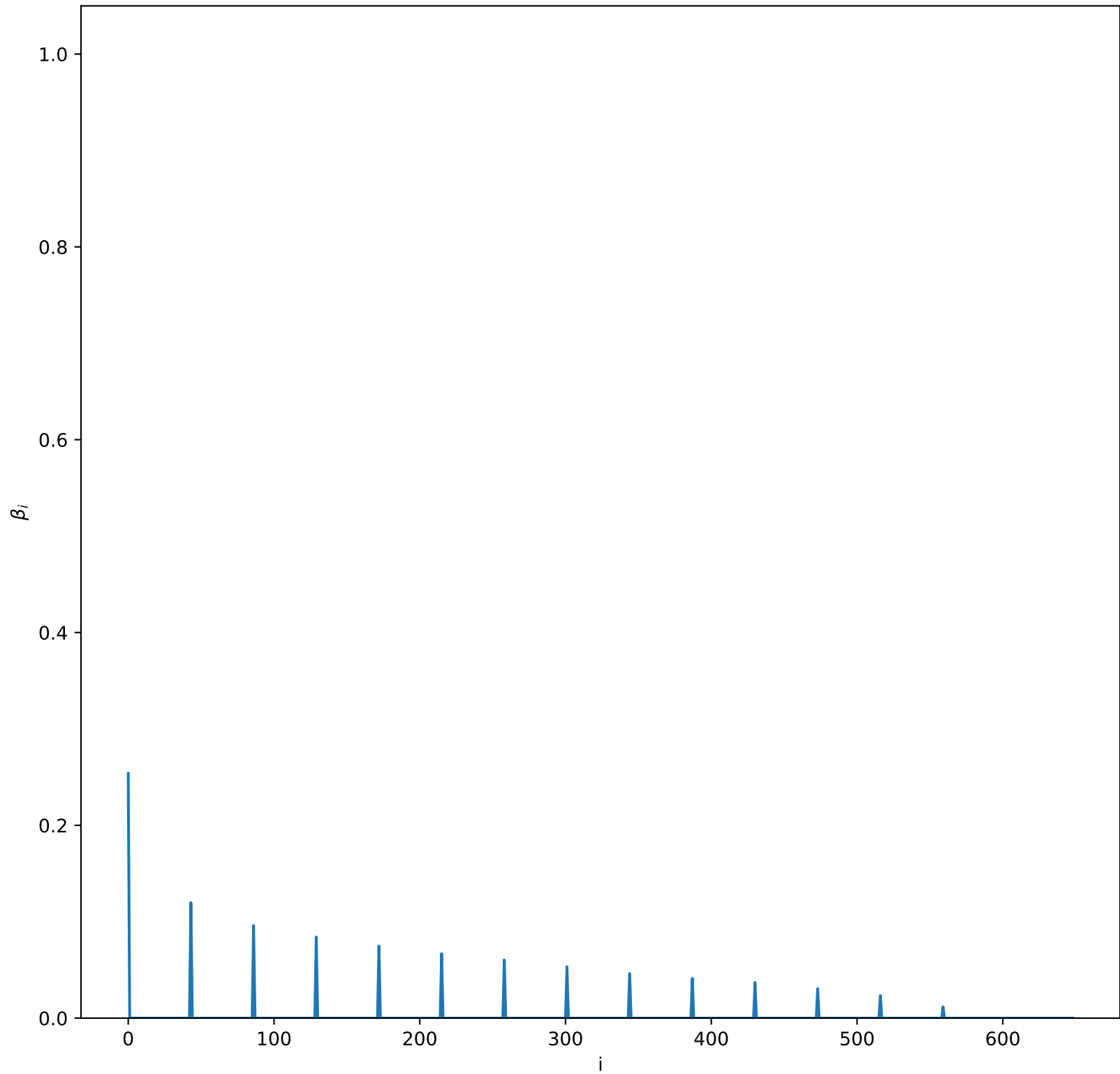
$\mu = 0.39$



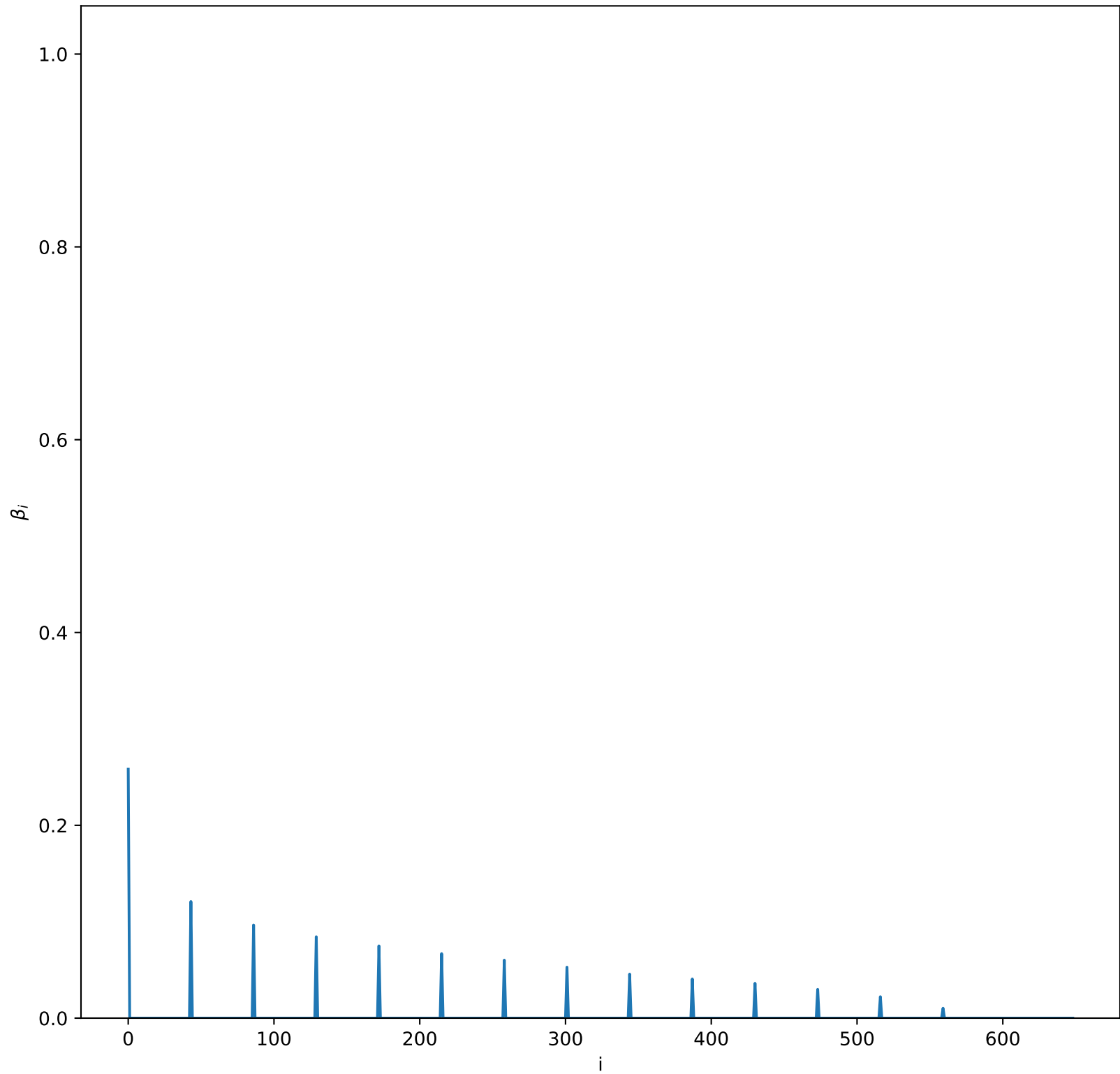
$\mu = 0.40$



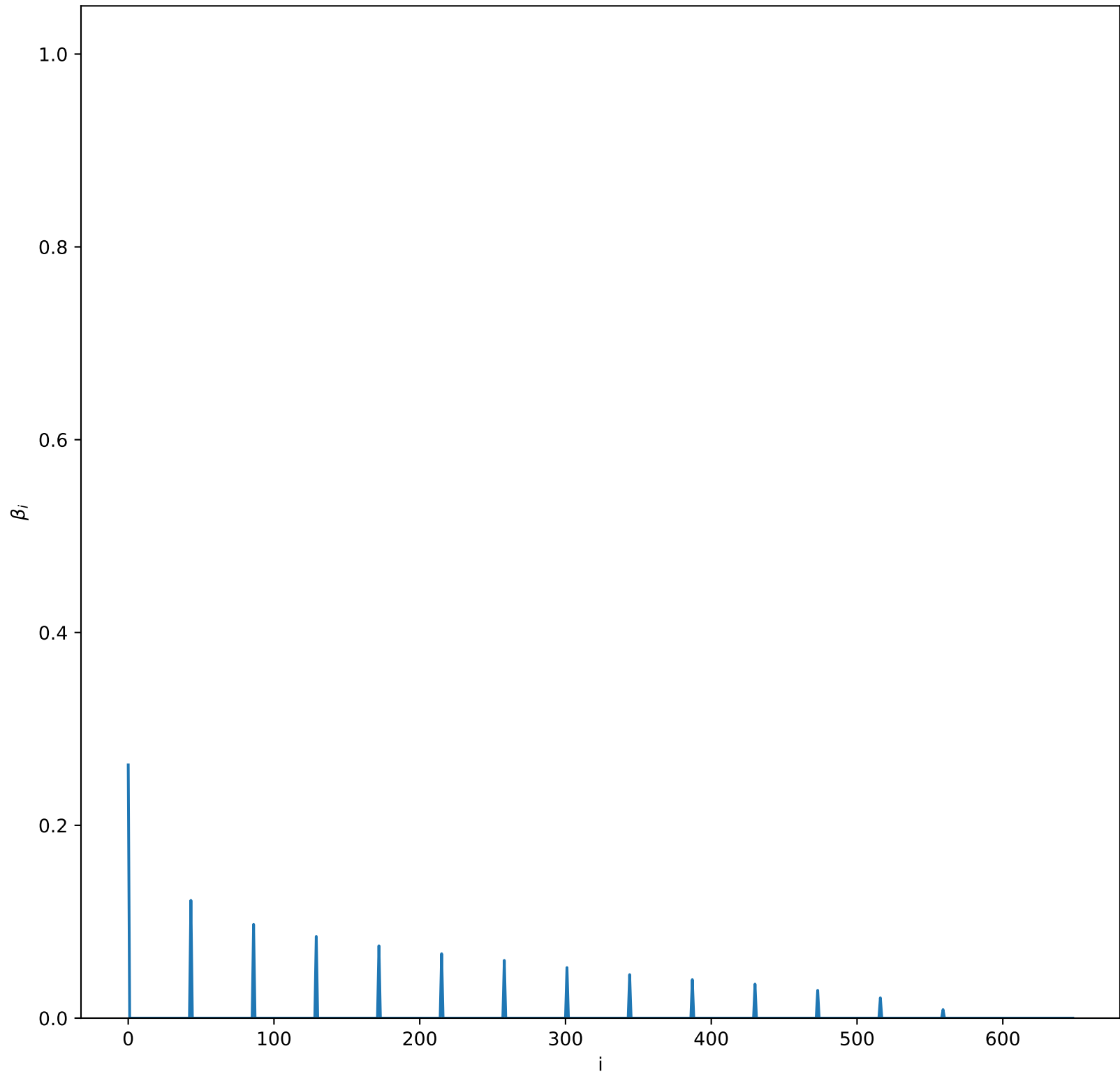
$\mu = 0.41$



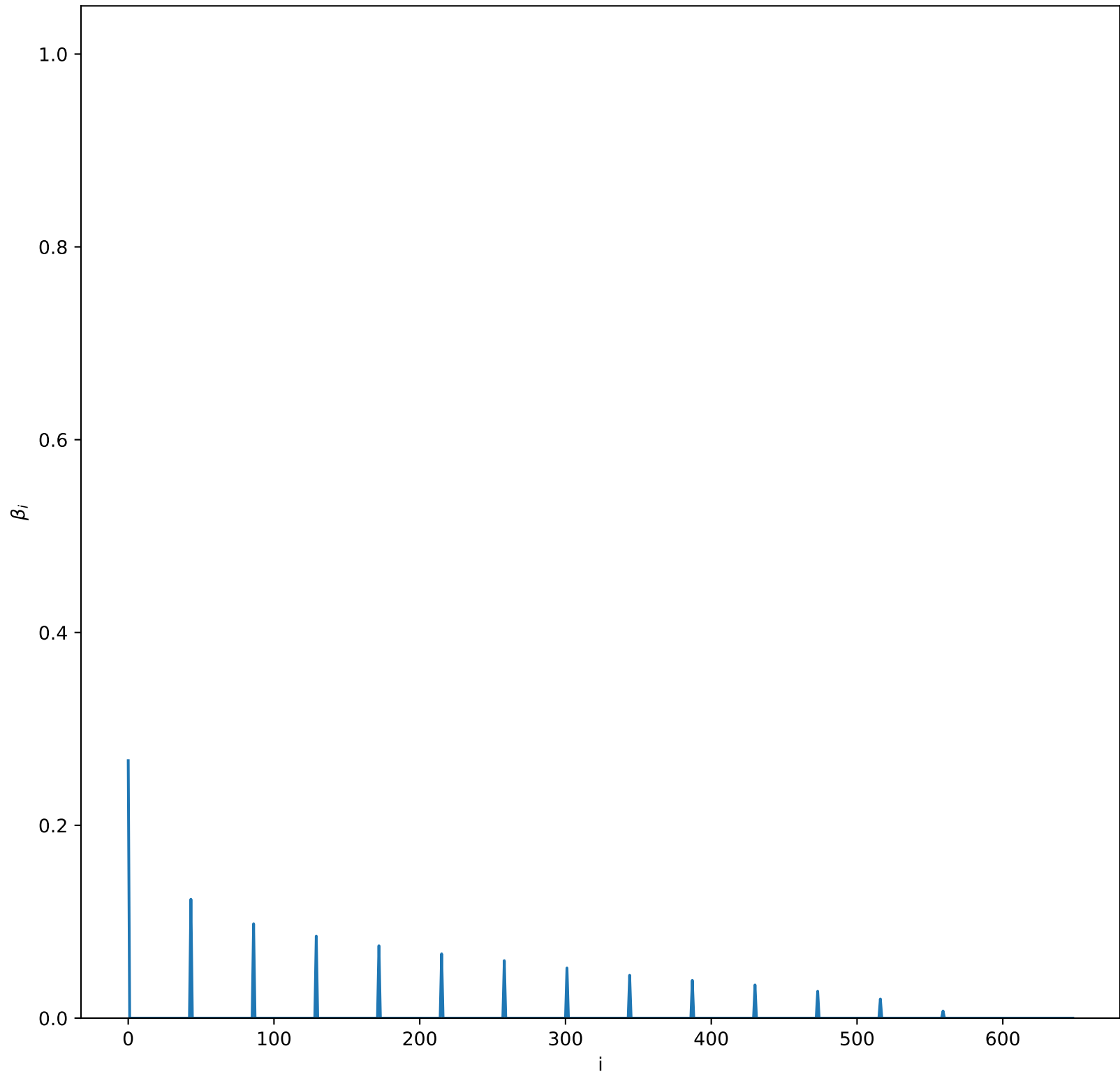
$\mu = 0.42$



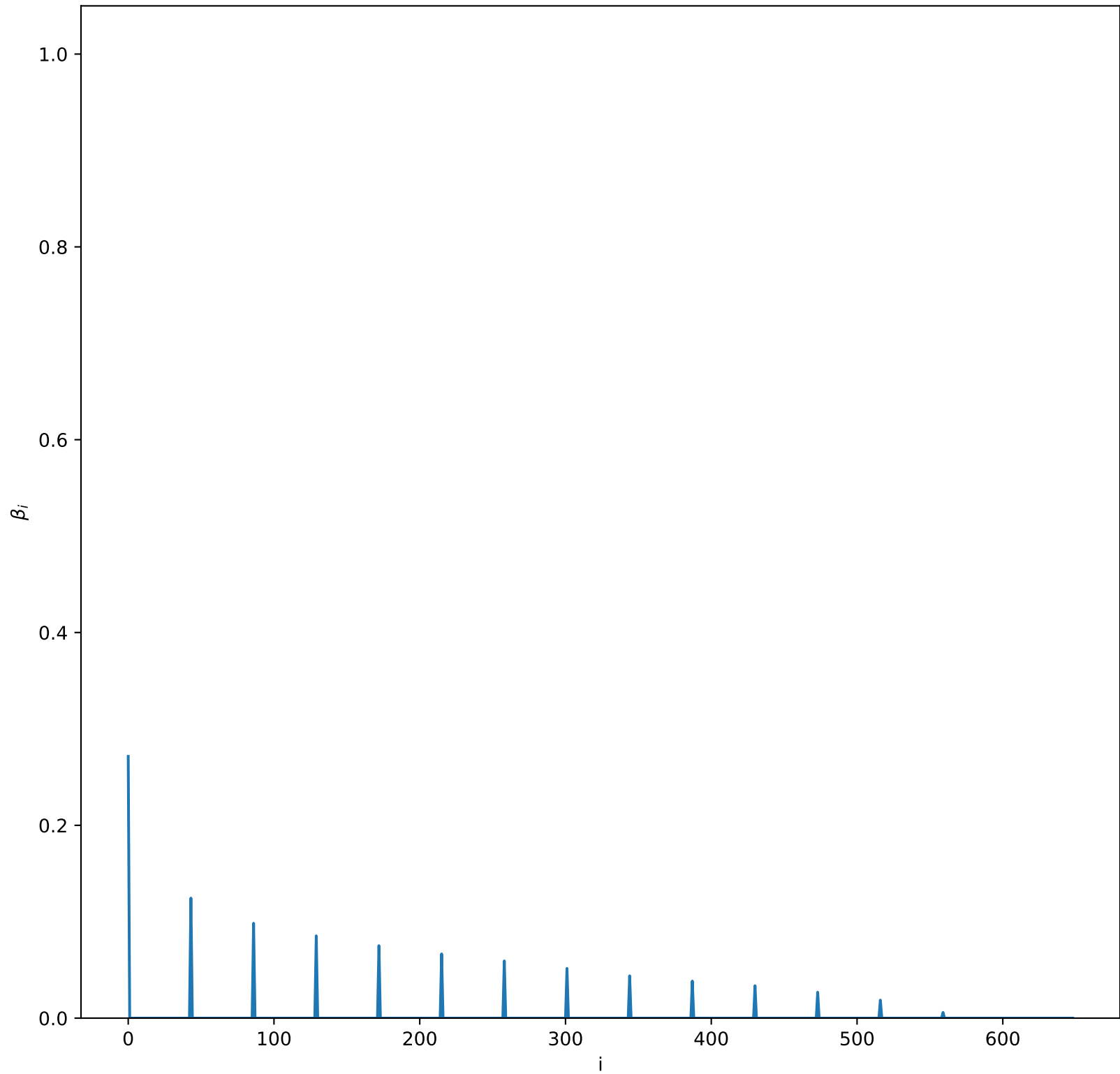
$\mu = 0.43$



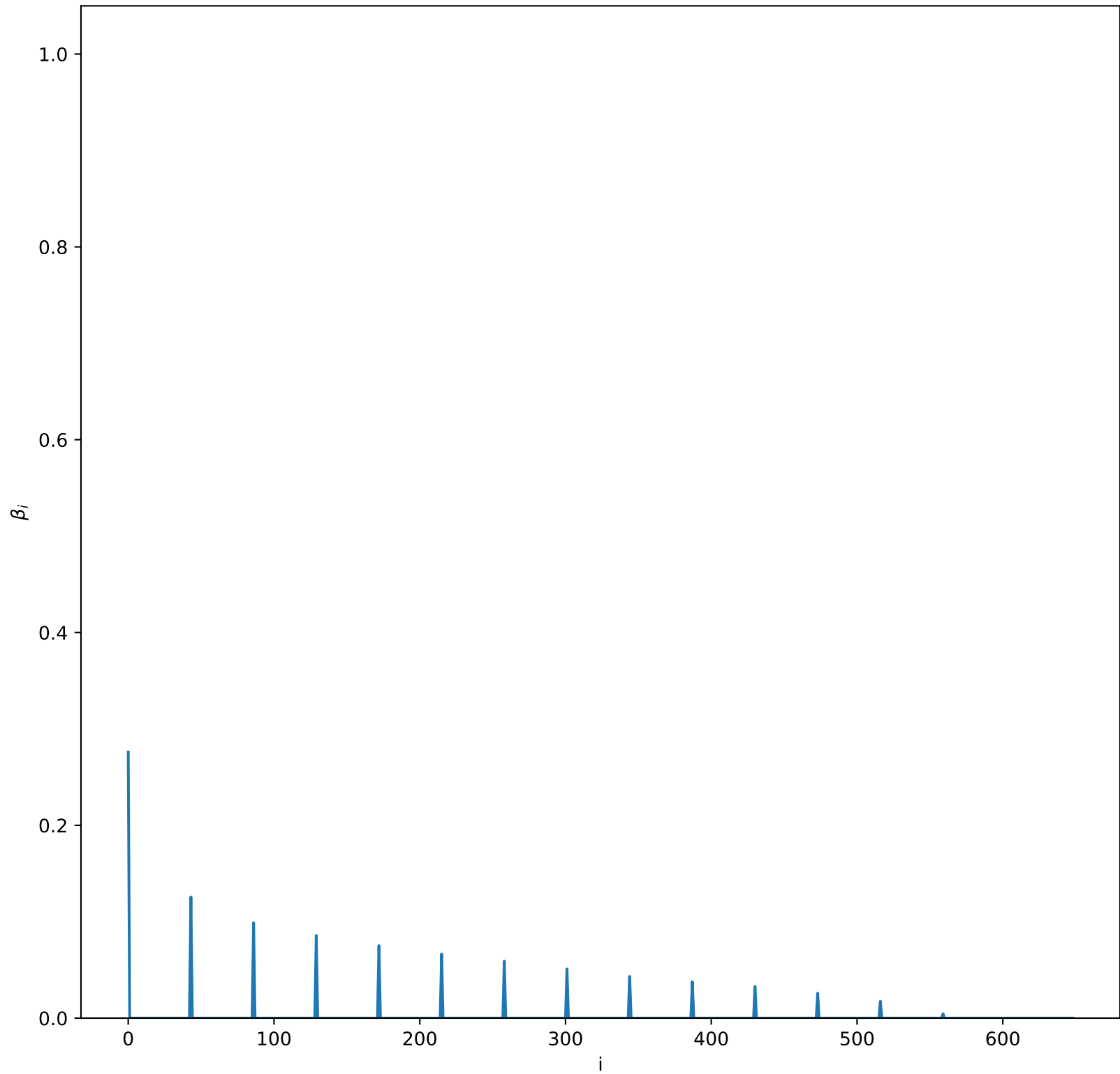
$\mu = 0.44$



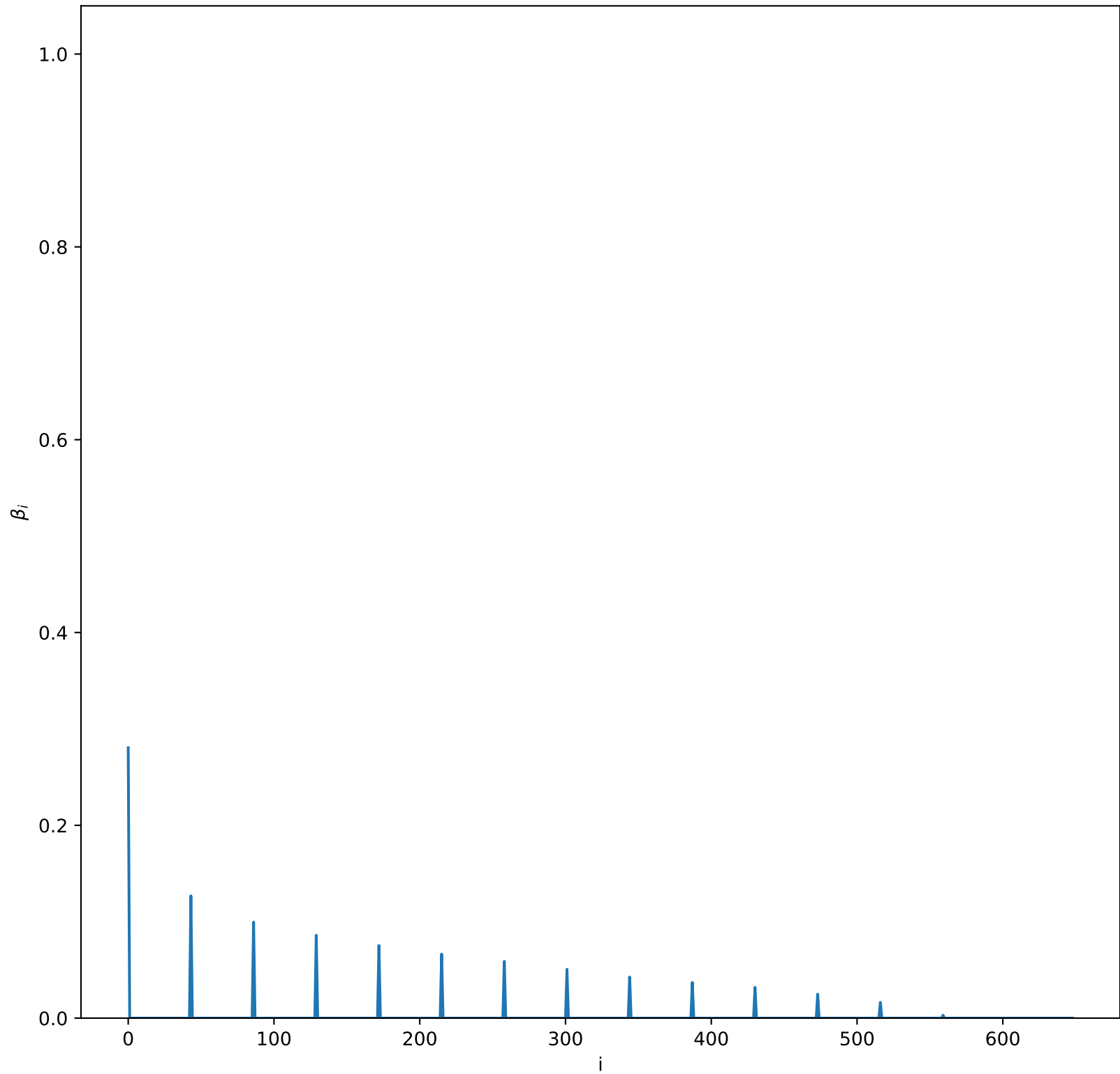
$\mu = 0.45$



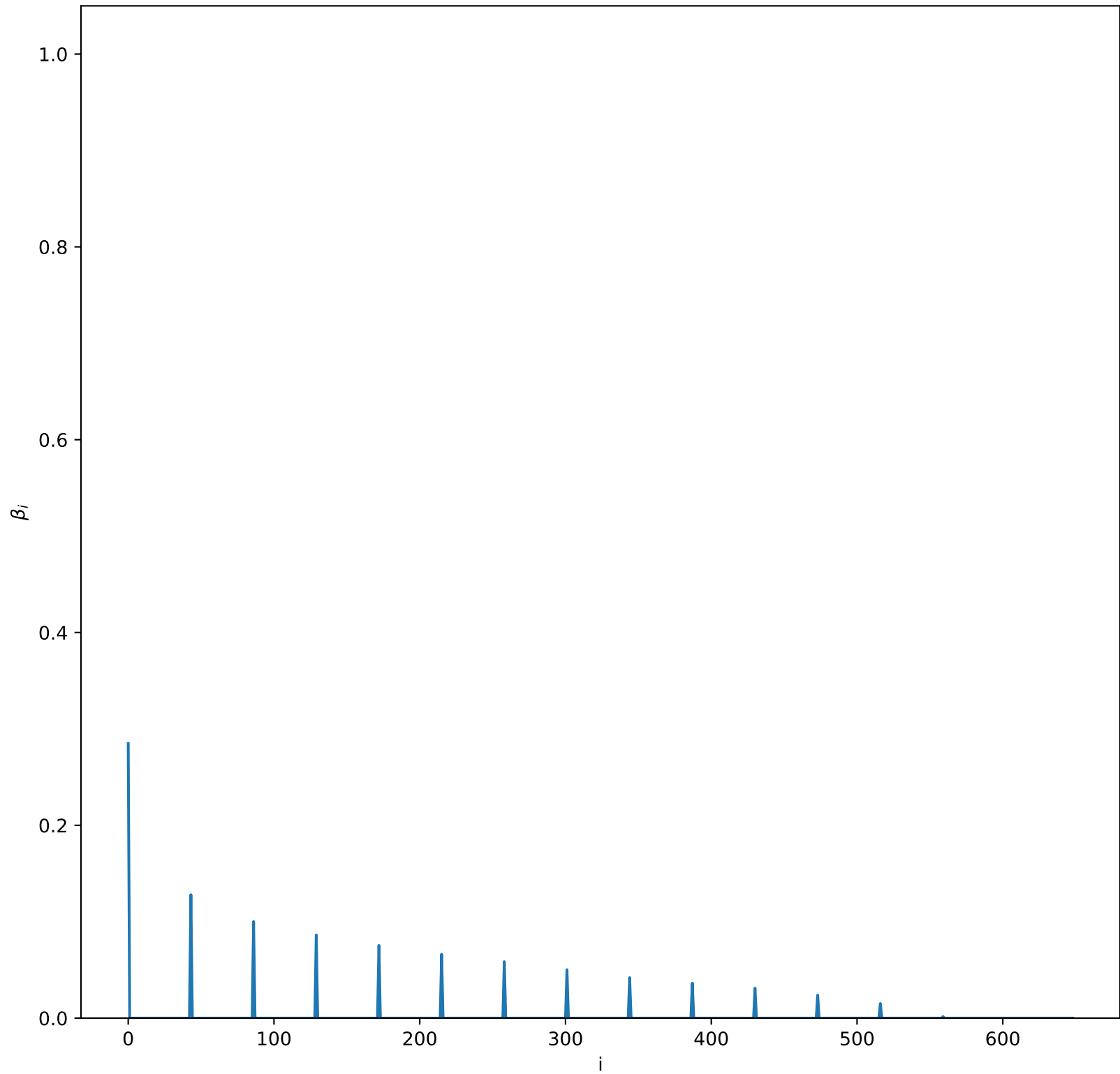
$\mu = 0.46$



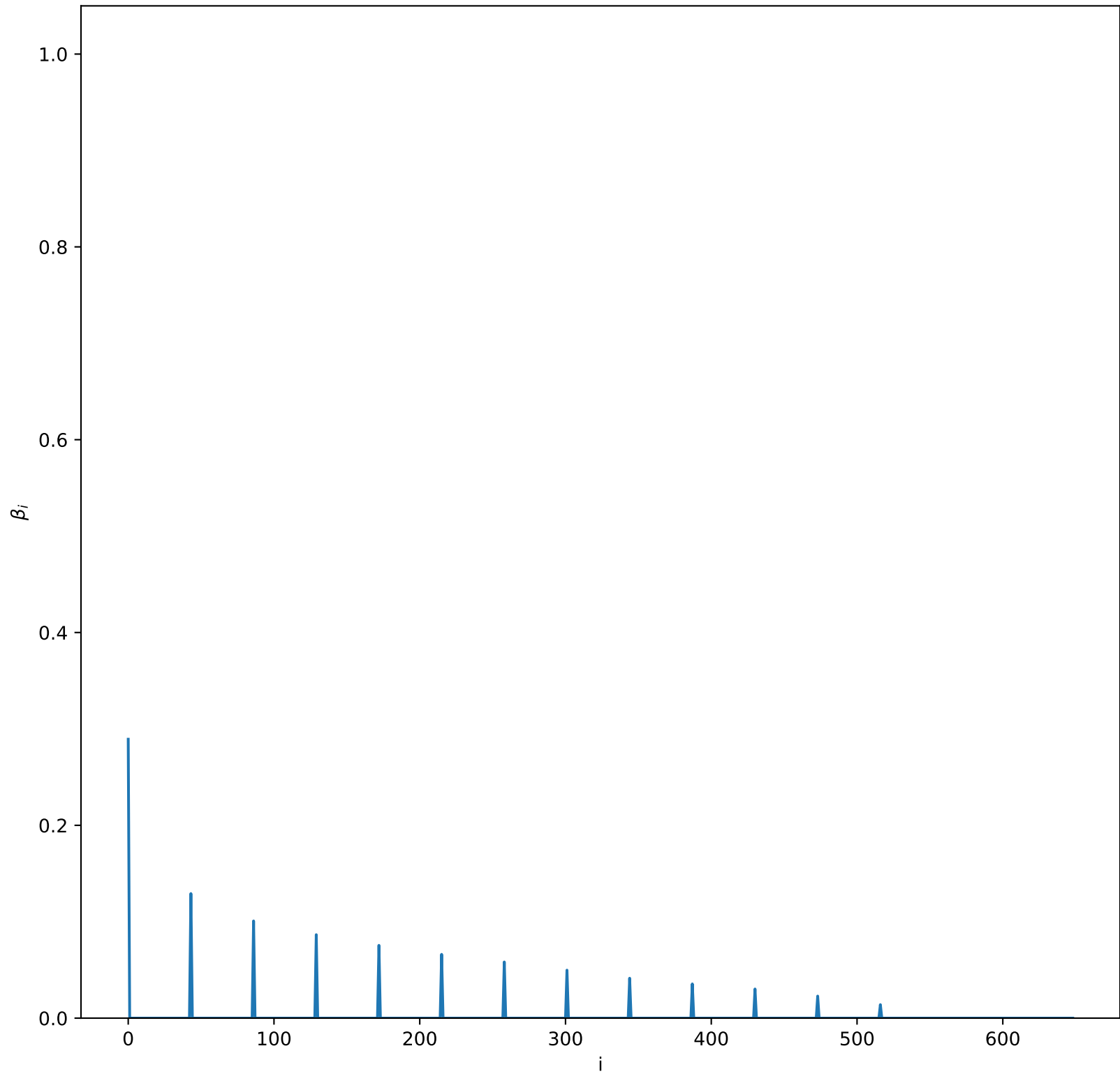
$\mu = 0.47$



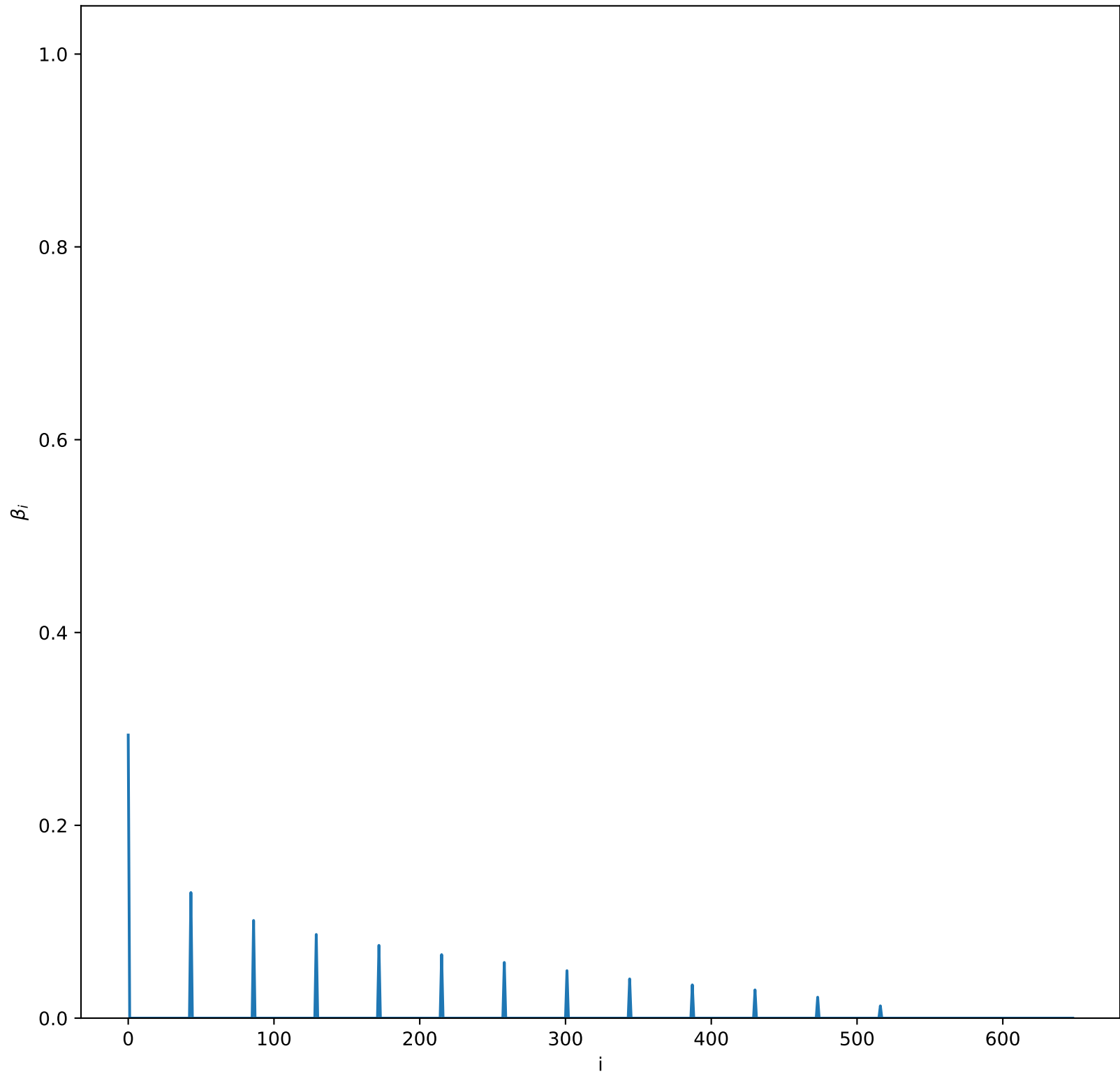
$\mu = 0.48$



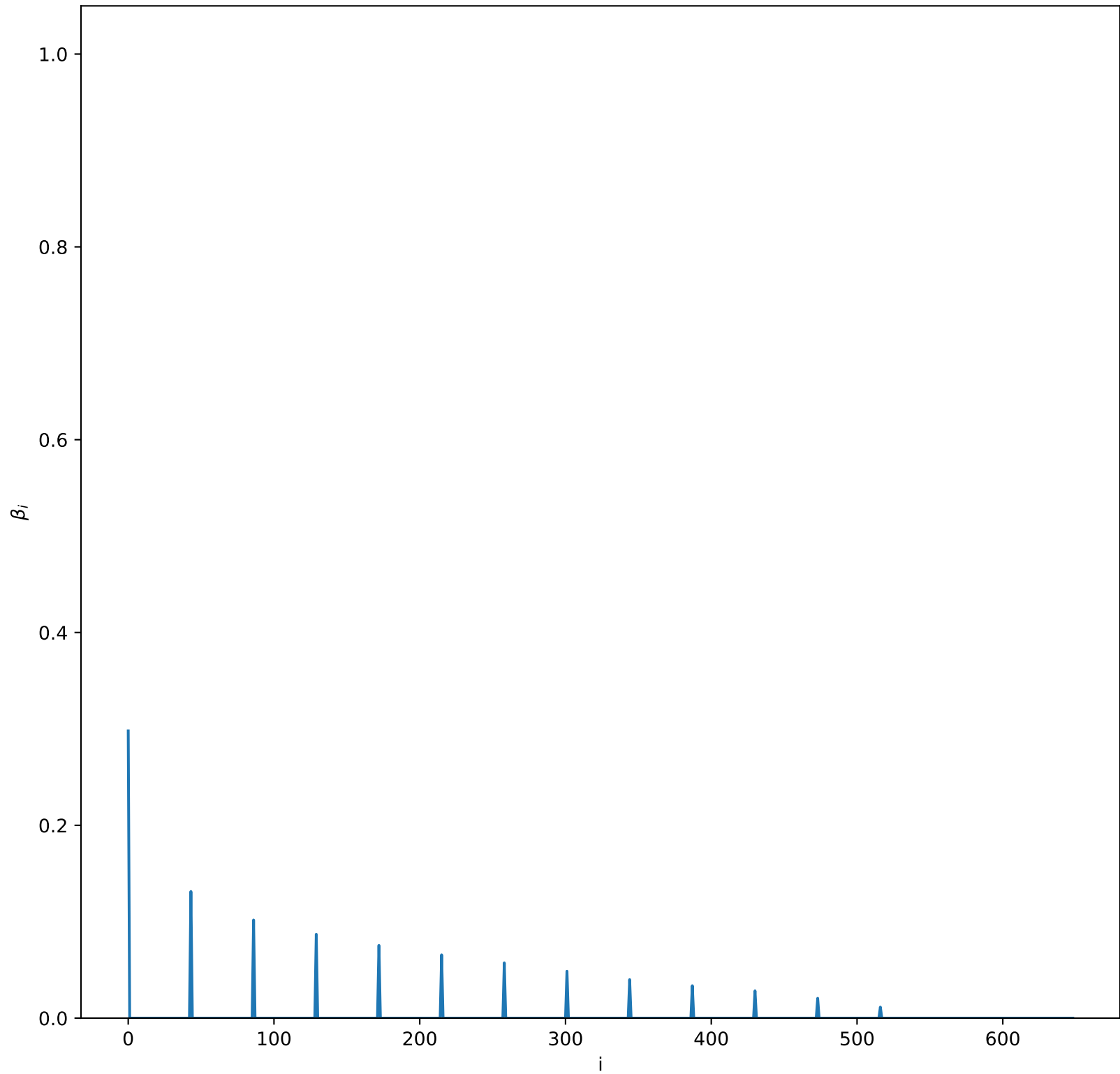
$\mu = 0.49$



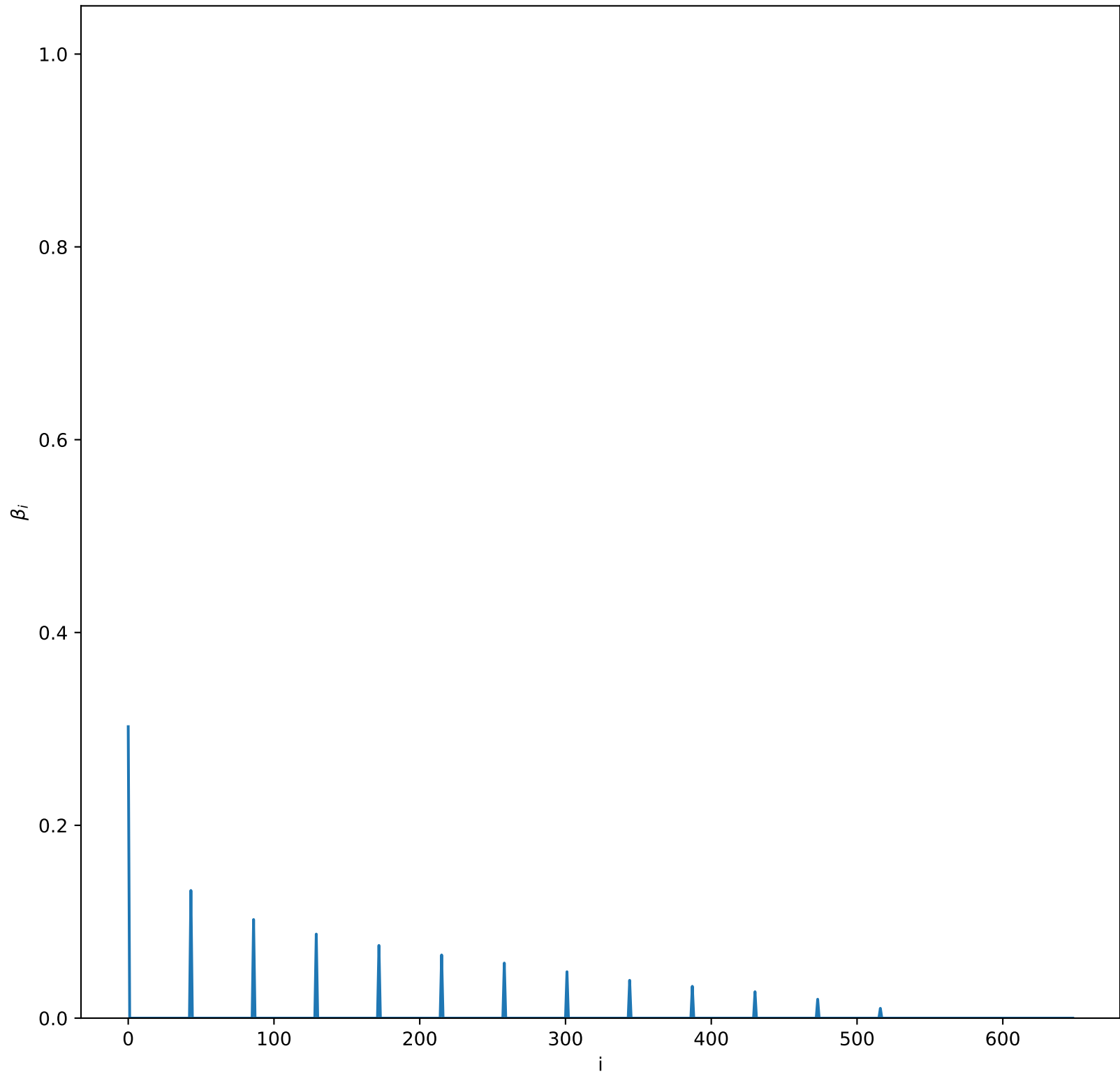
$\mu = 0.50$



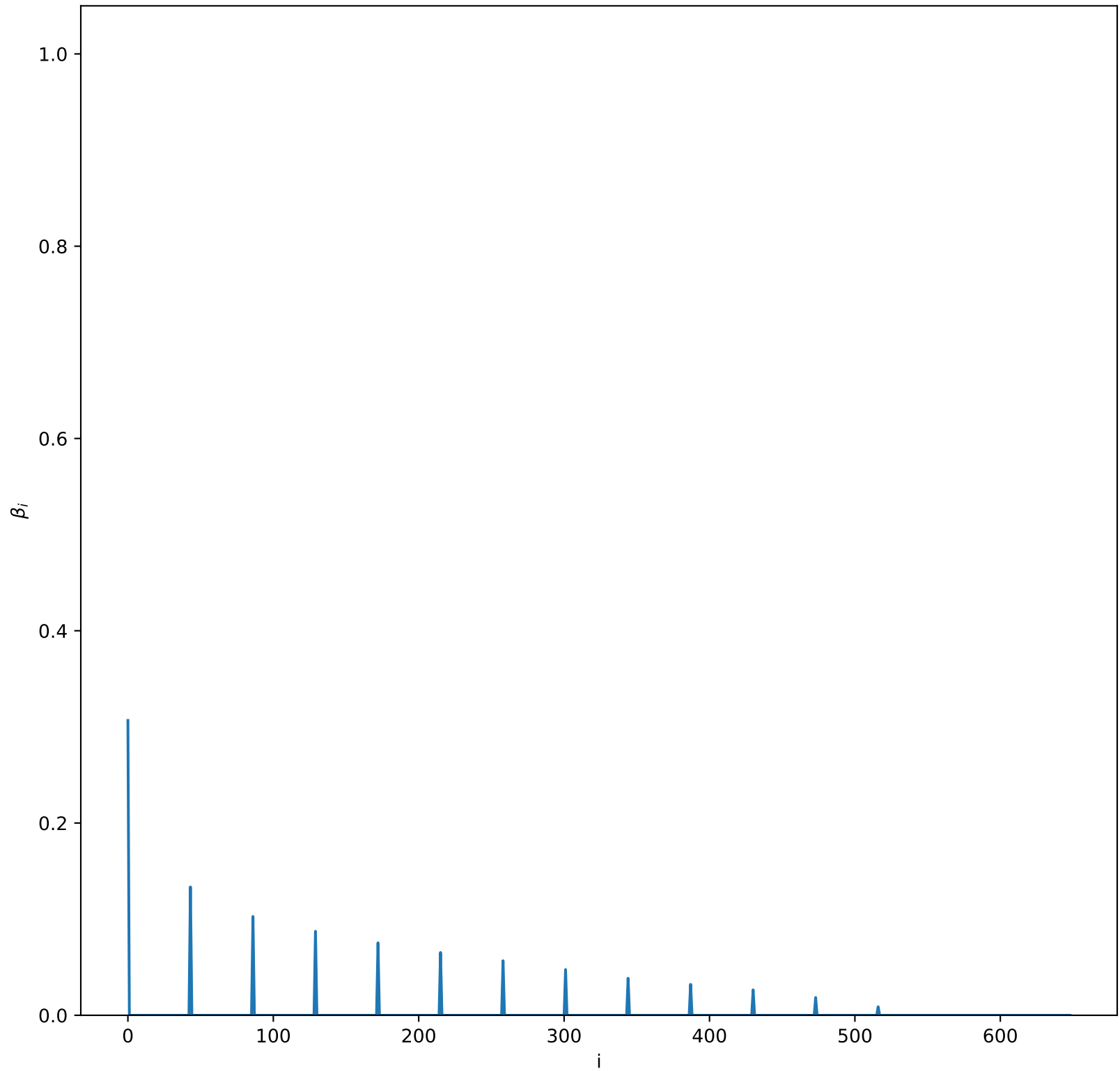
$\mu = 0.51$



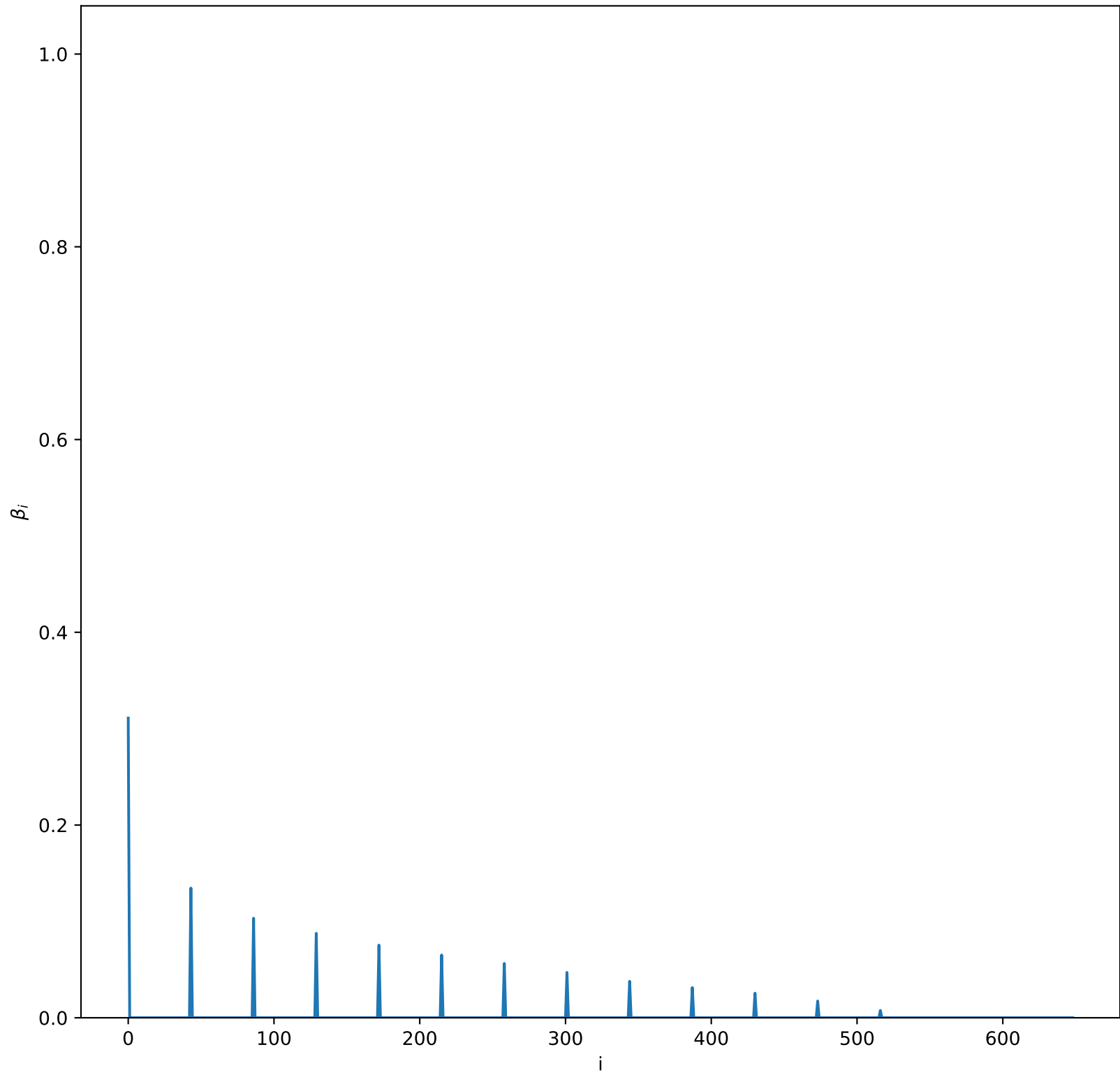
$\mu = 0.52$



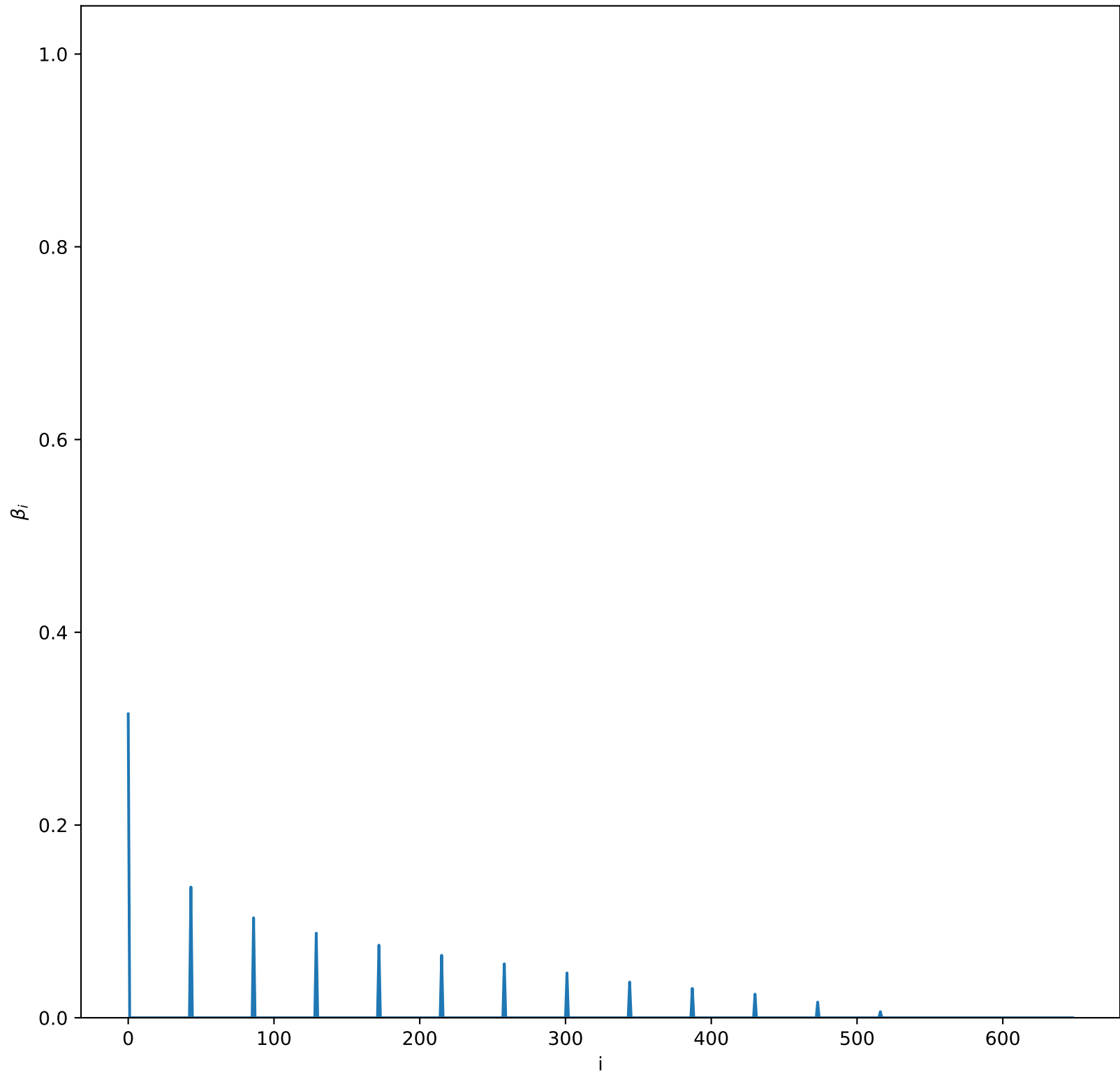
$\mu = 0.53$



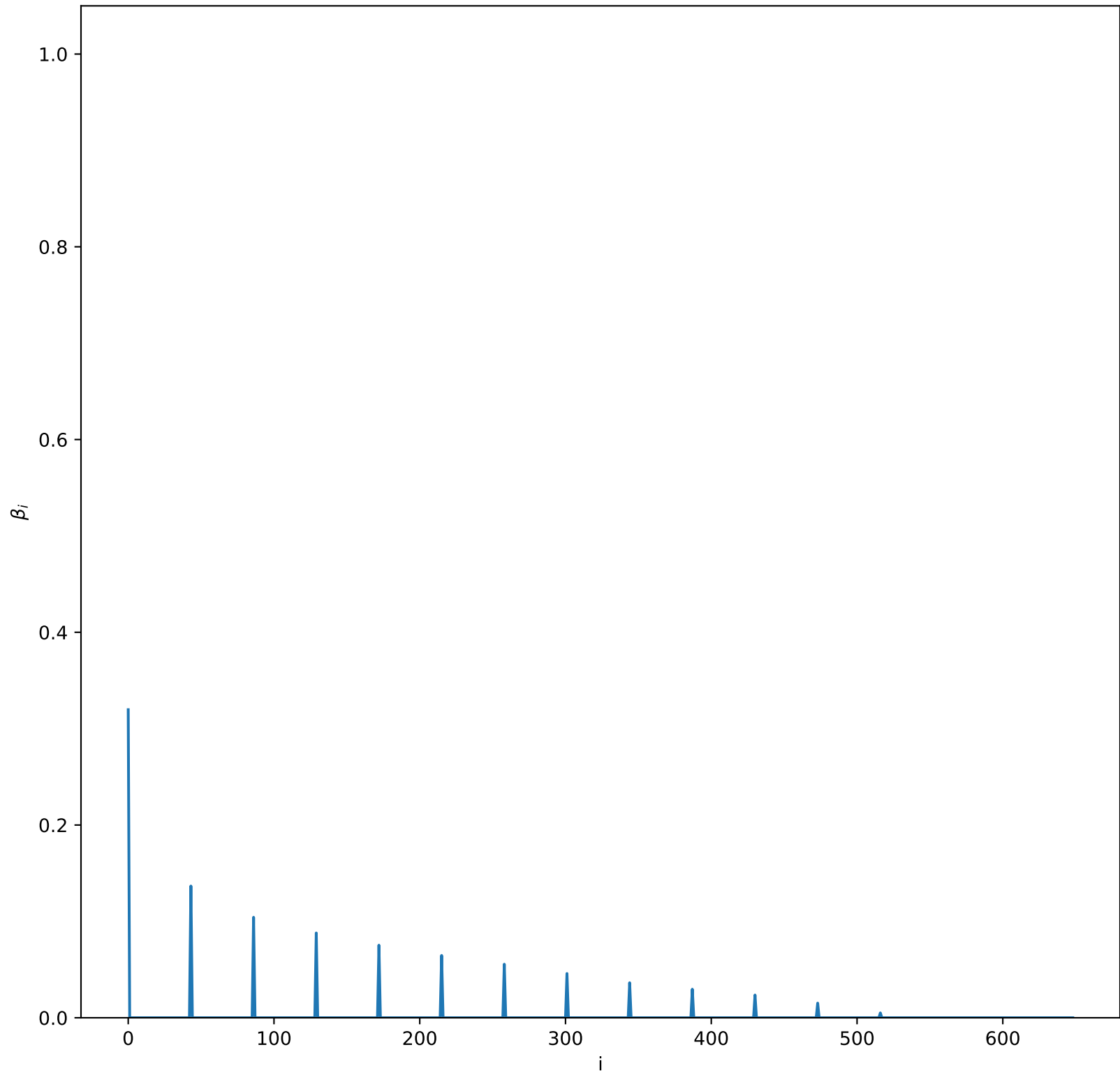
$\mu = 0.54$



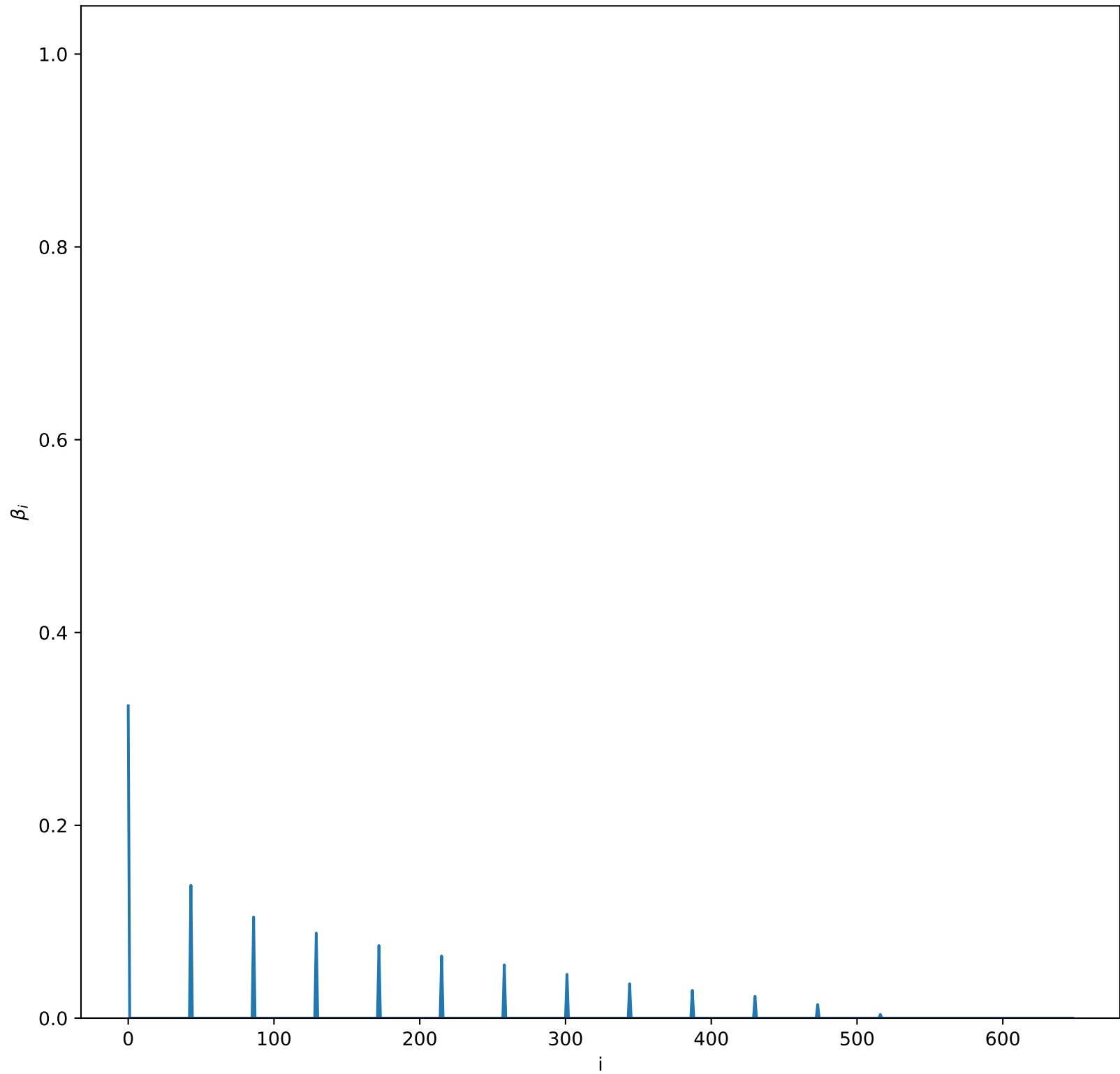
$\mu = 0.55$



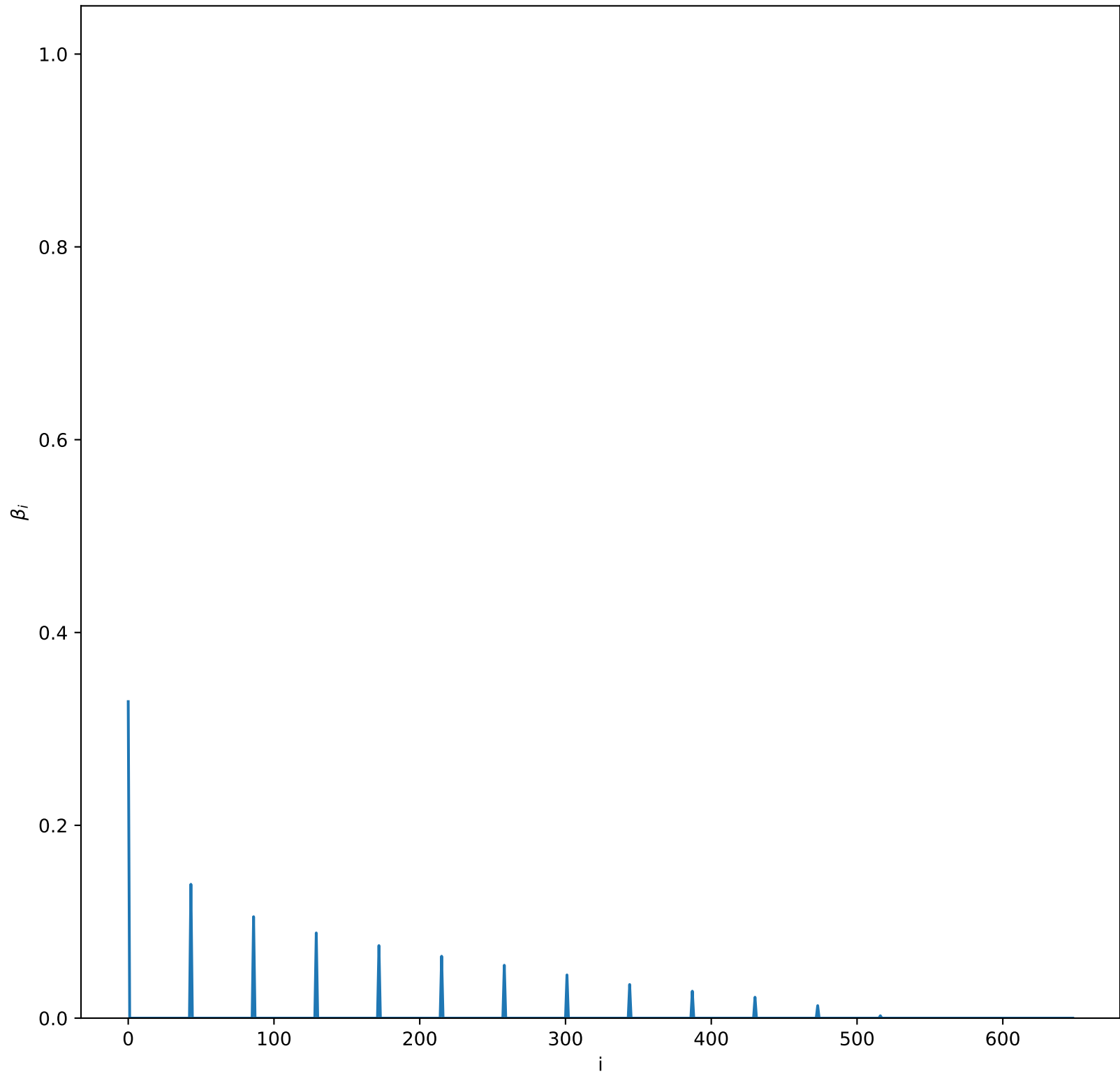
$\mu = 0.56$



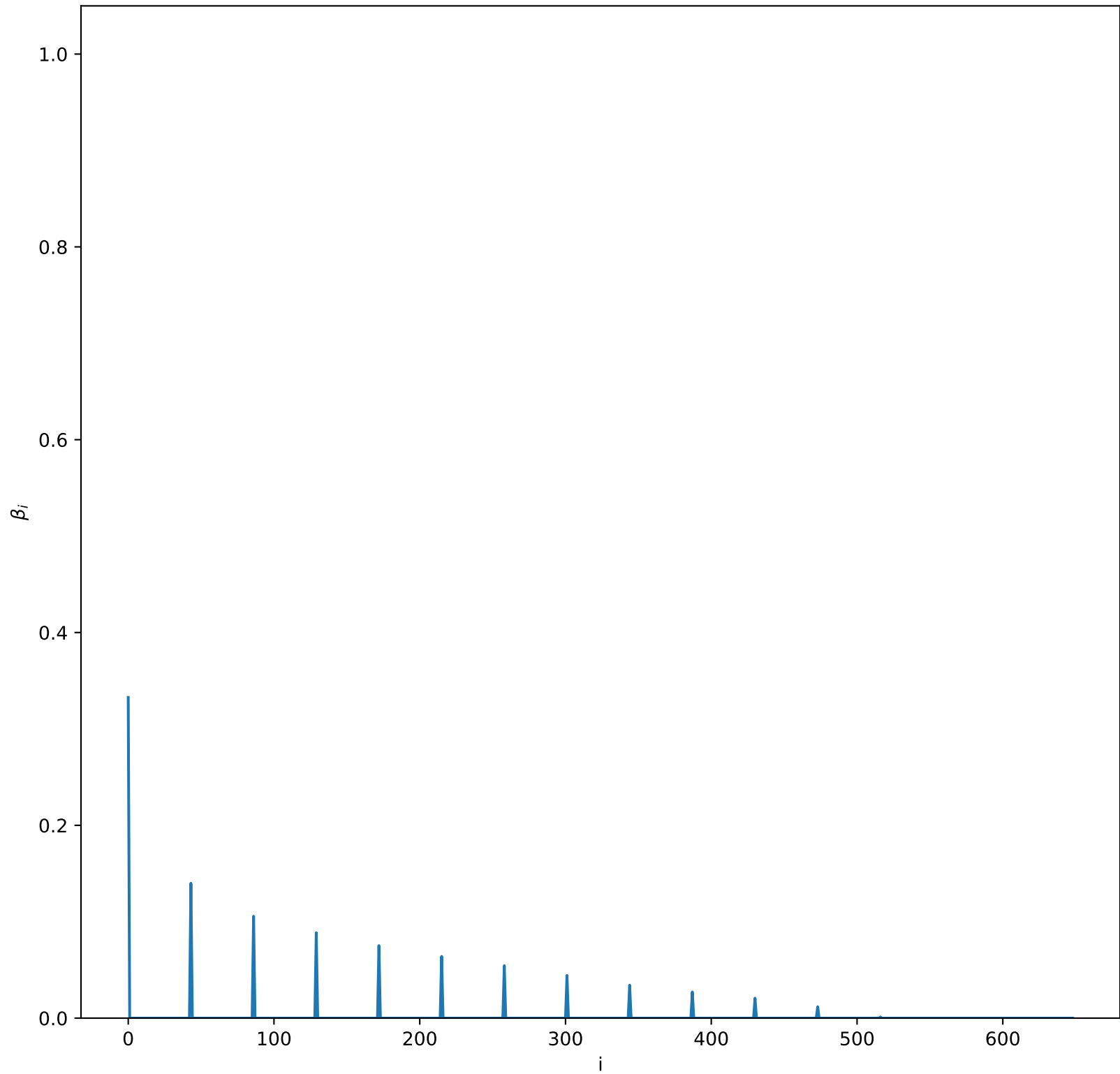
$\mu = 0.57$



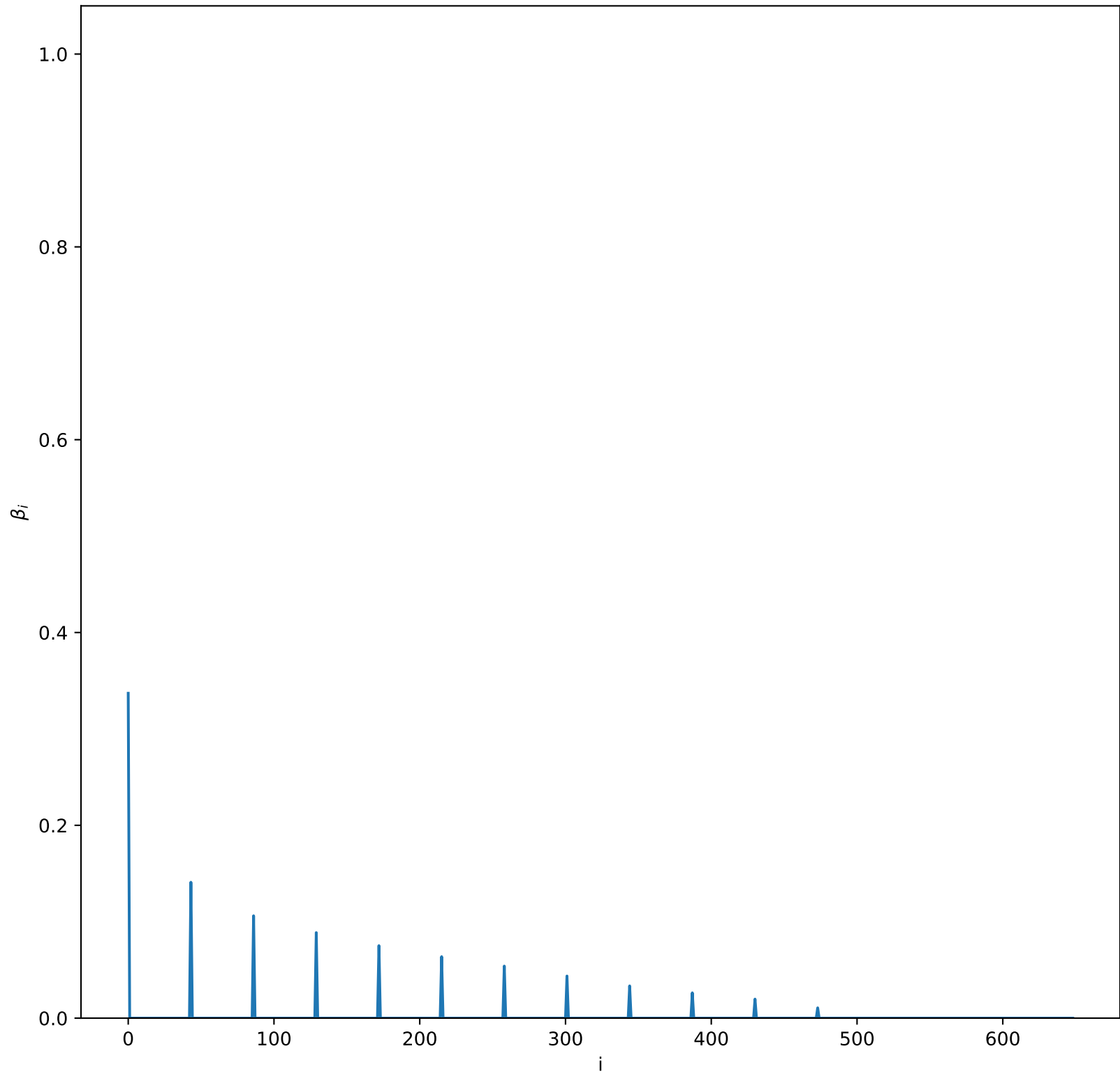
$\mu = 0.58$



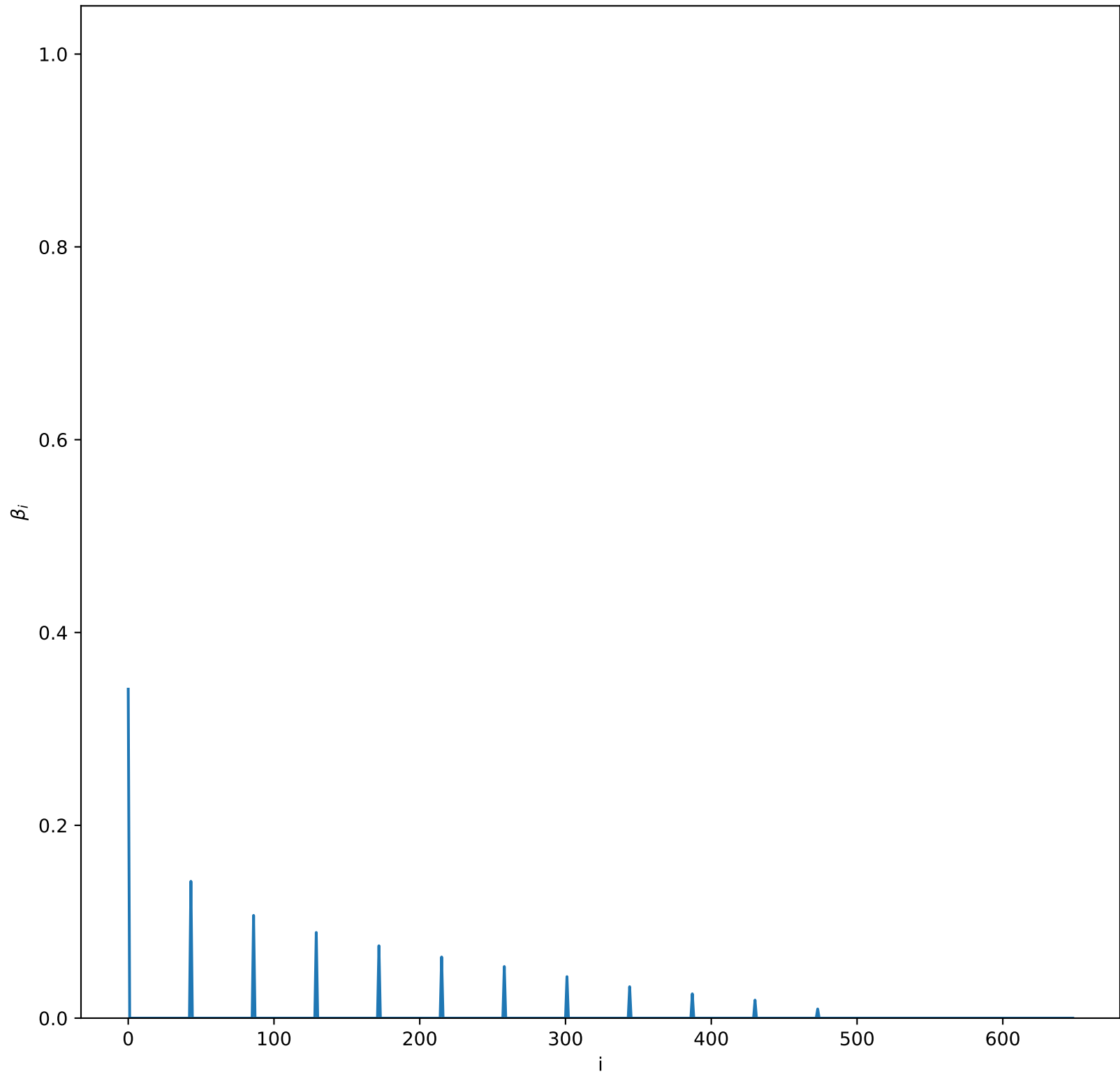
$\mu = 0.59$



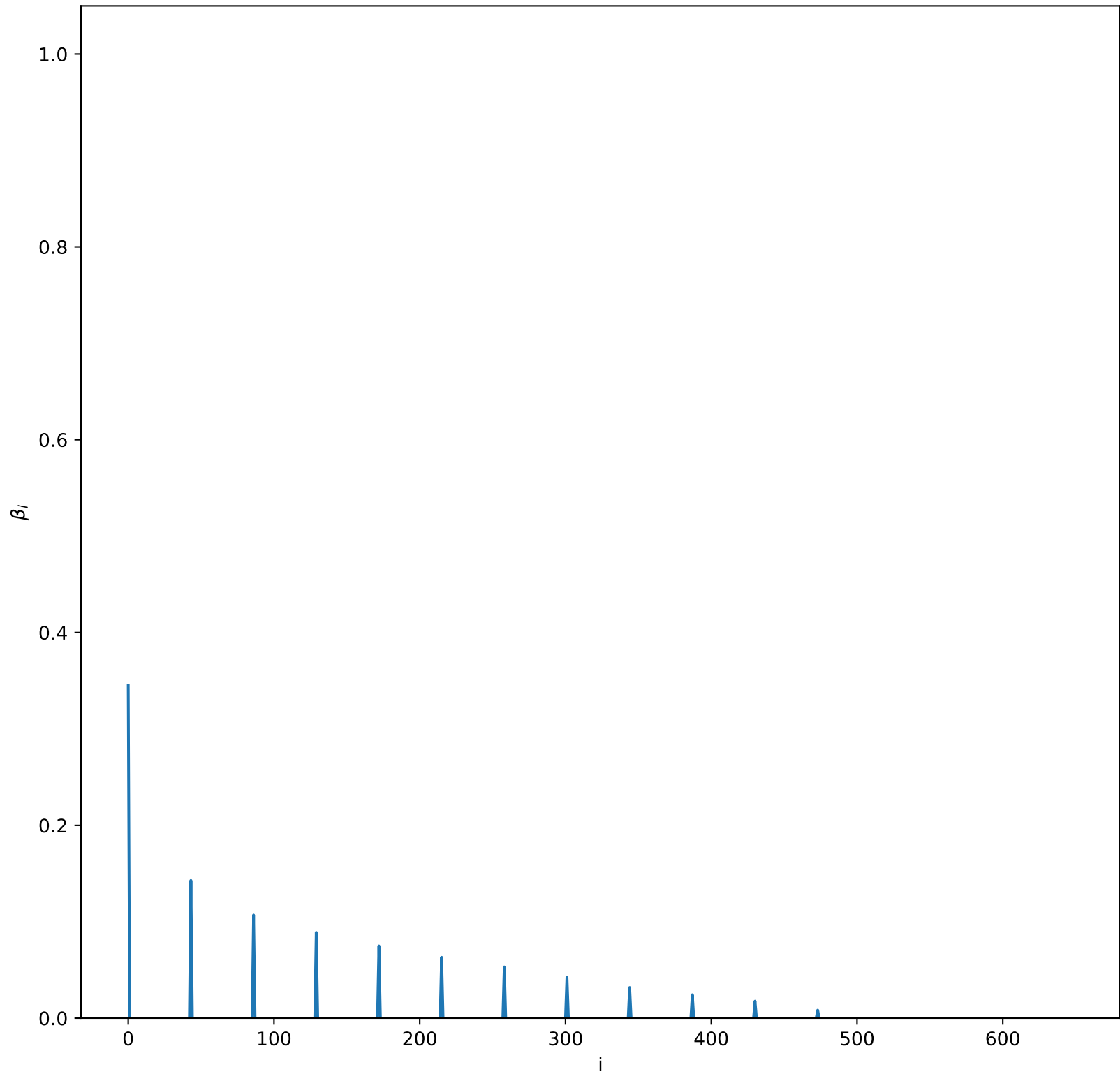
$\mu = 0.60$



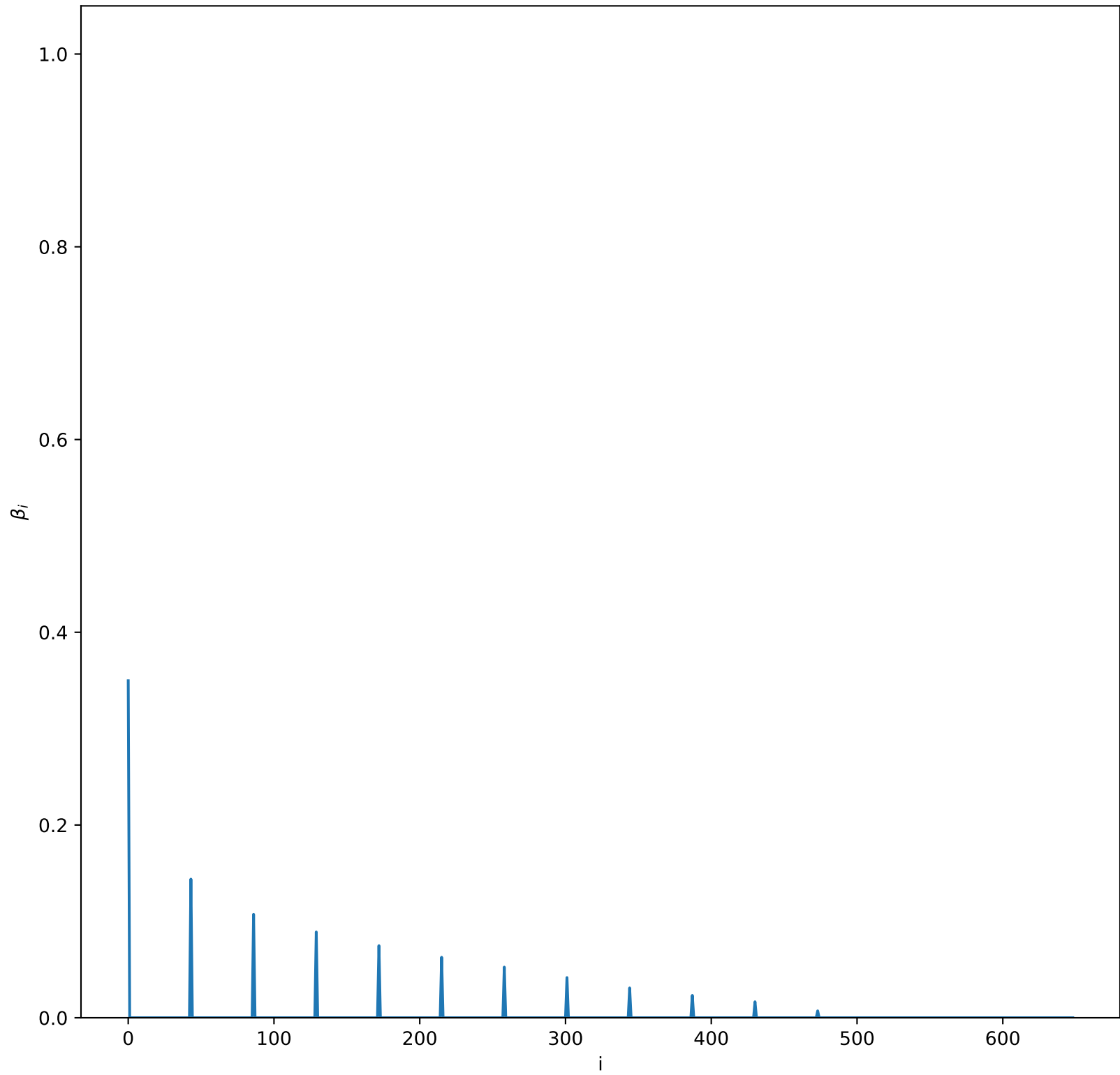
$\mu = 0.61$



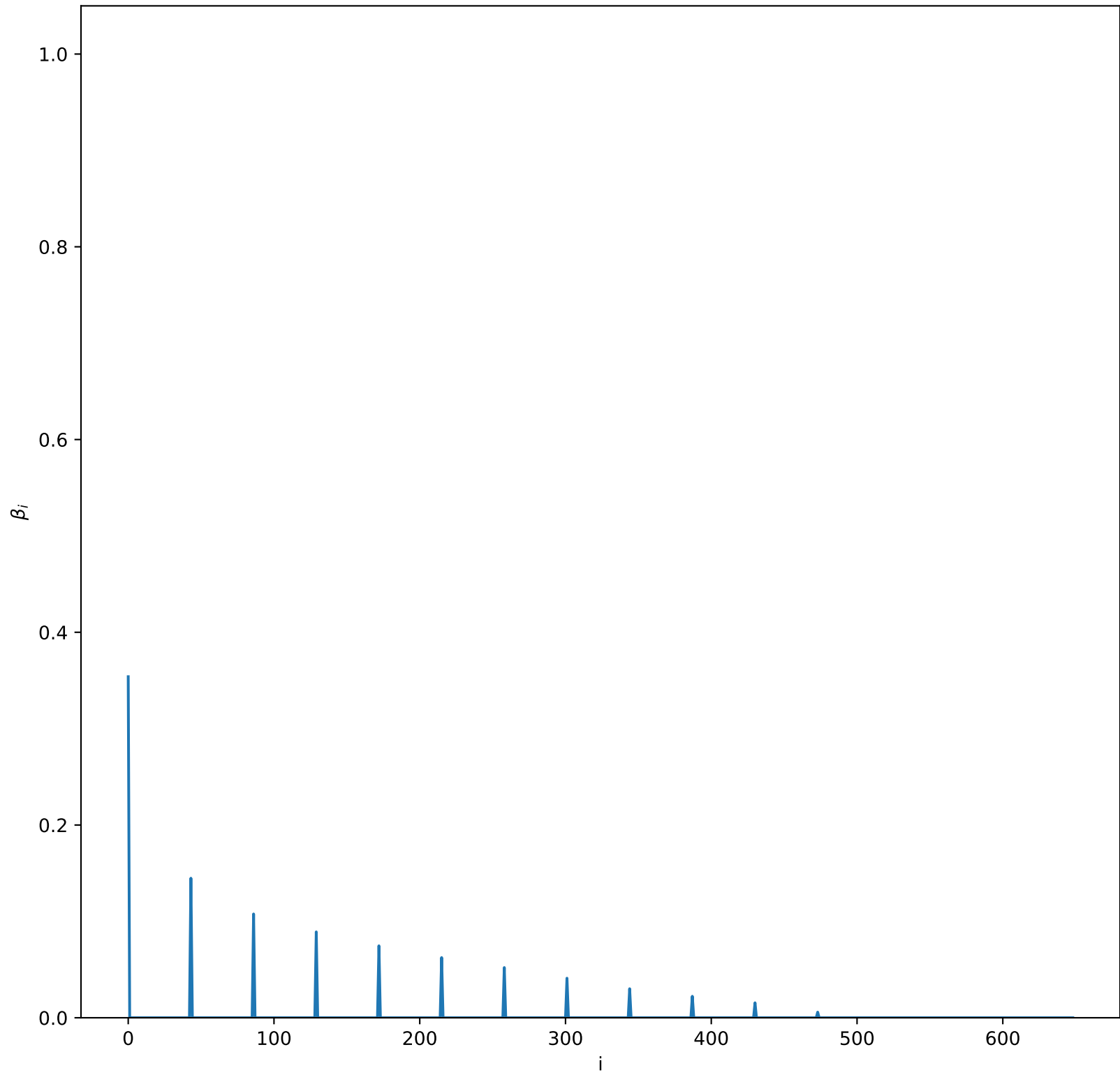
$\mu = 0.62$



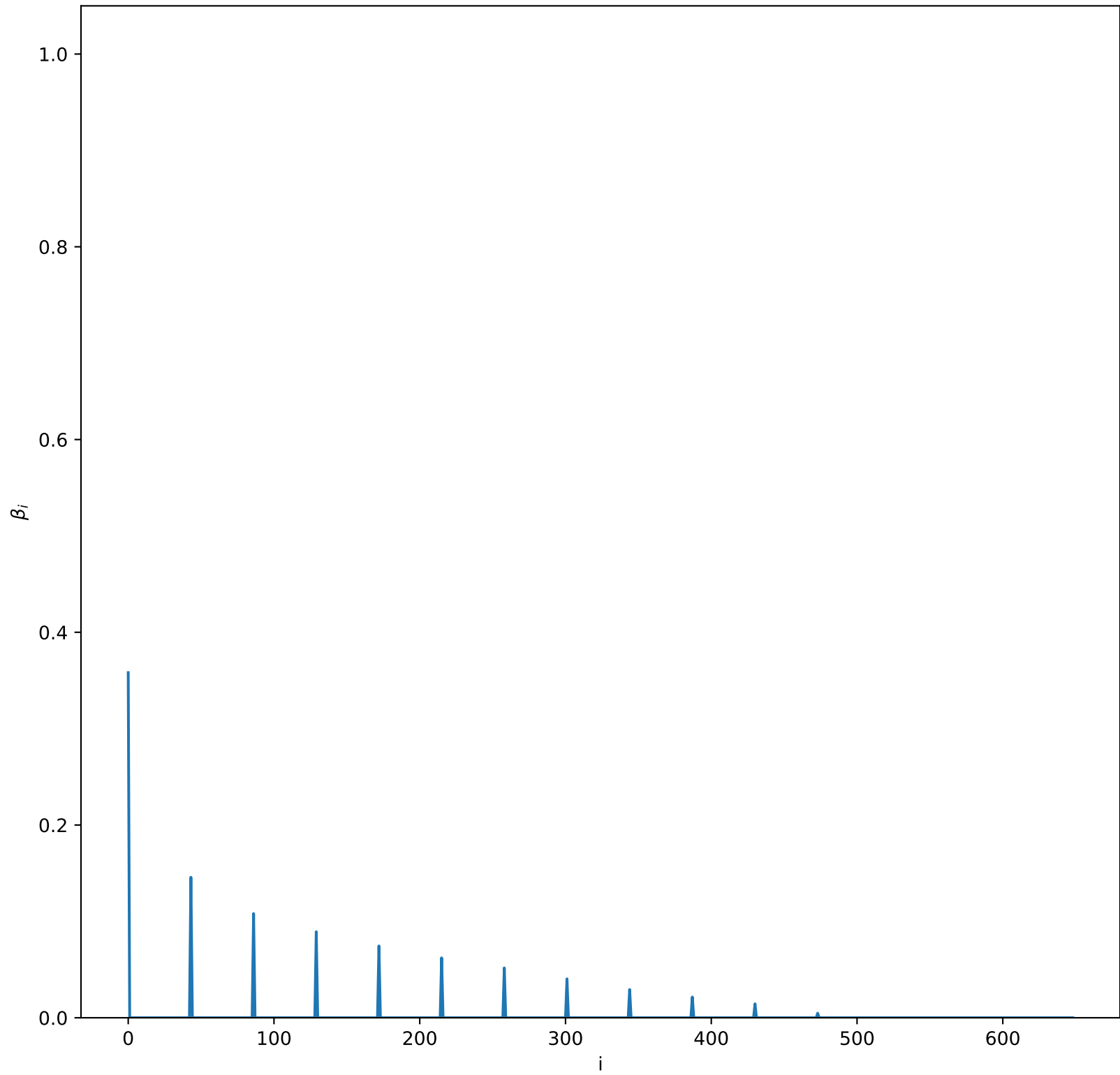
$\mu = 0.63$



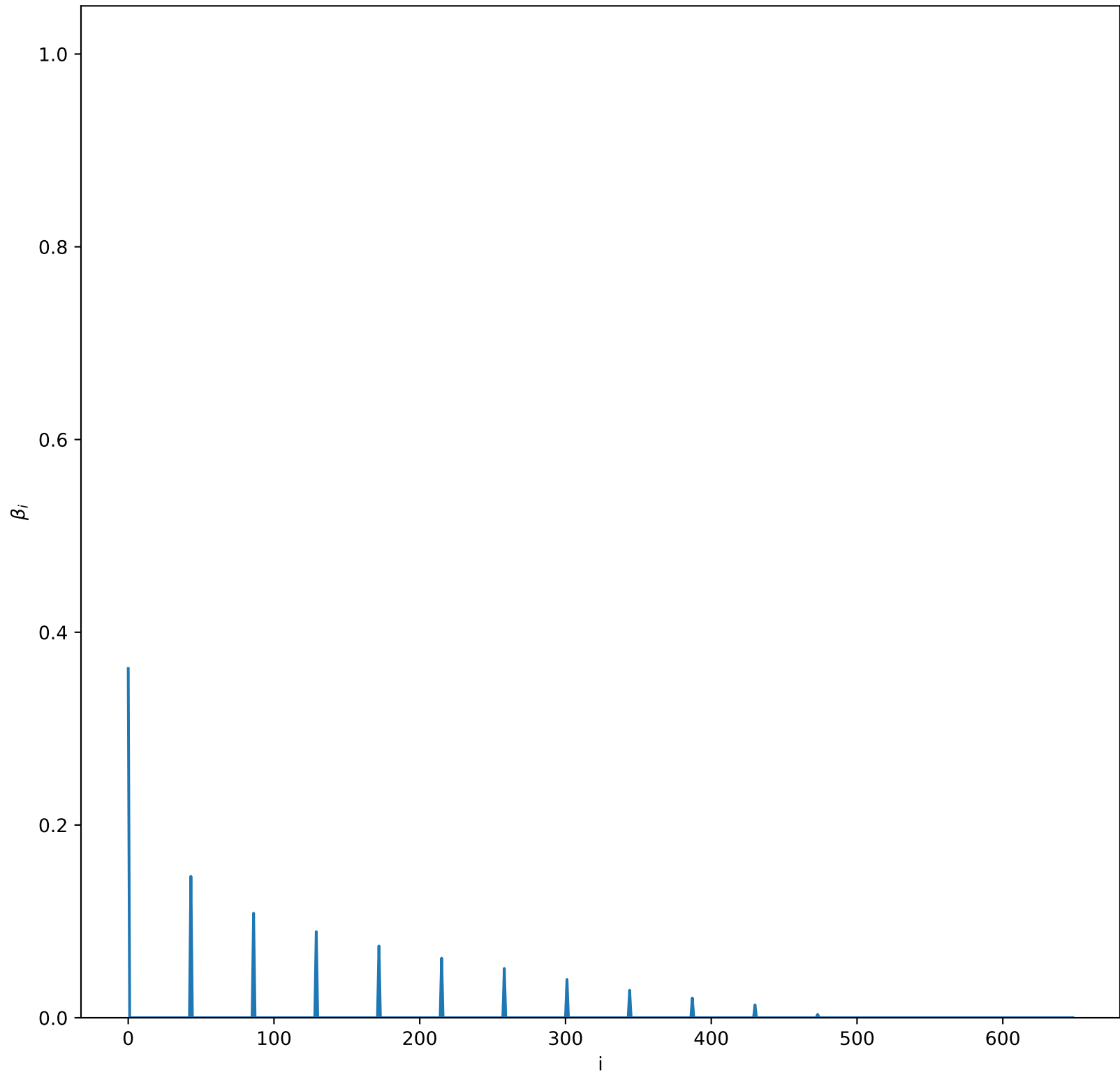
$\mu = 0.64$



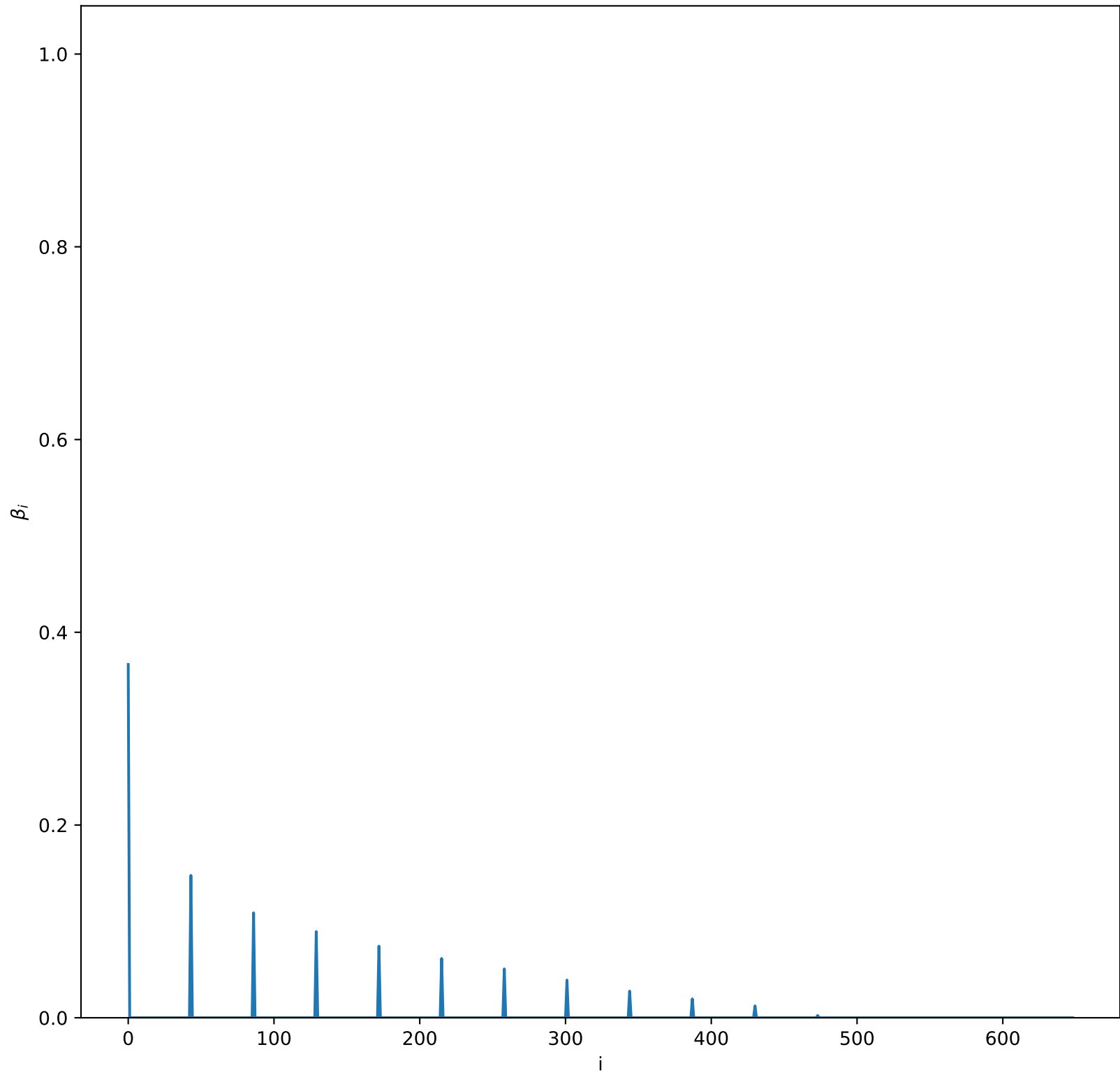
$\mu = 0.65$



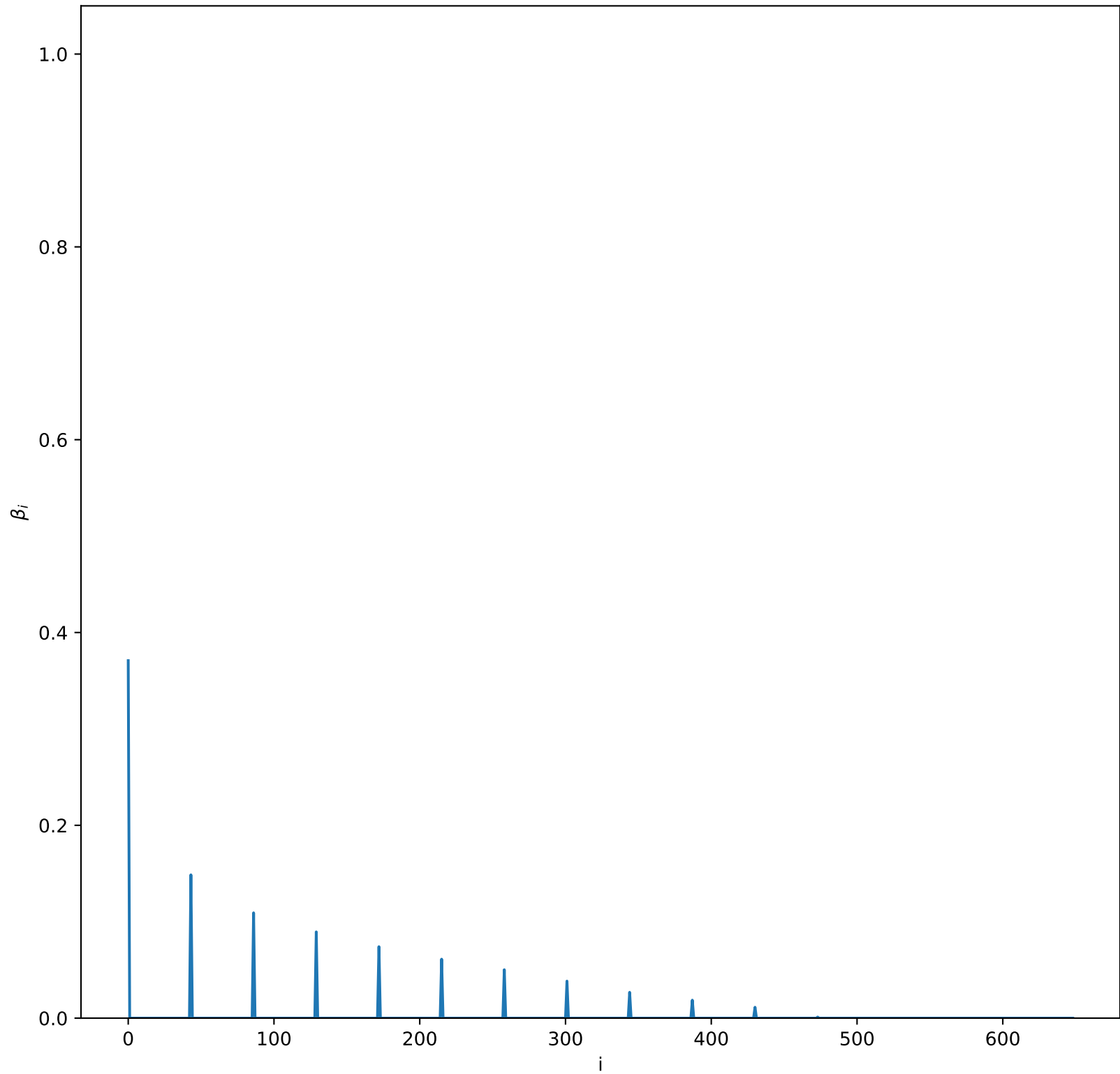
$\mu = 0.66$



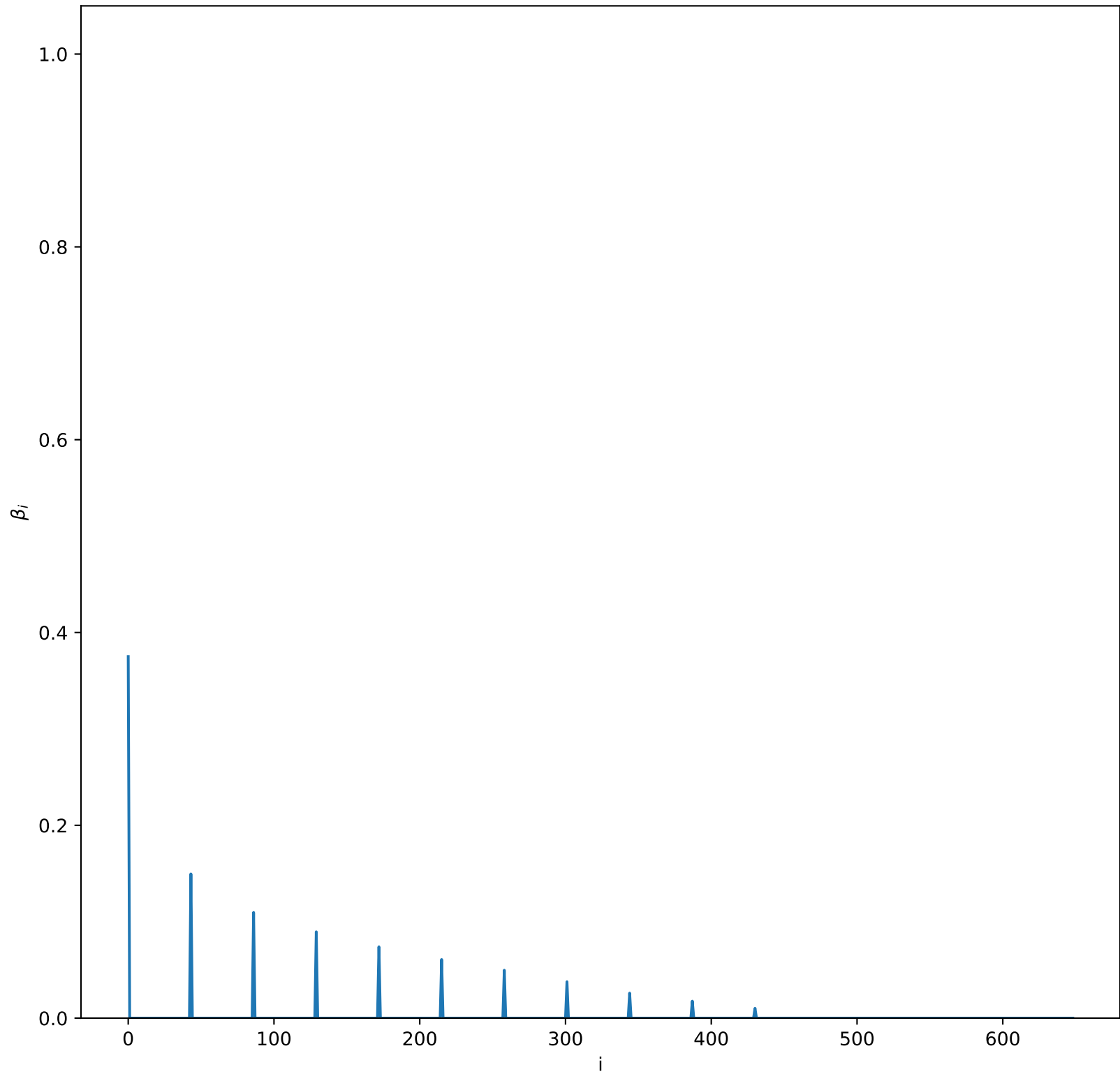
$\mu = 0.67$



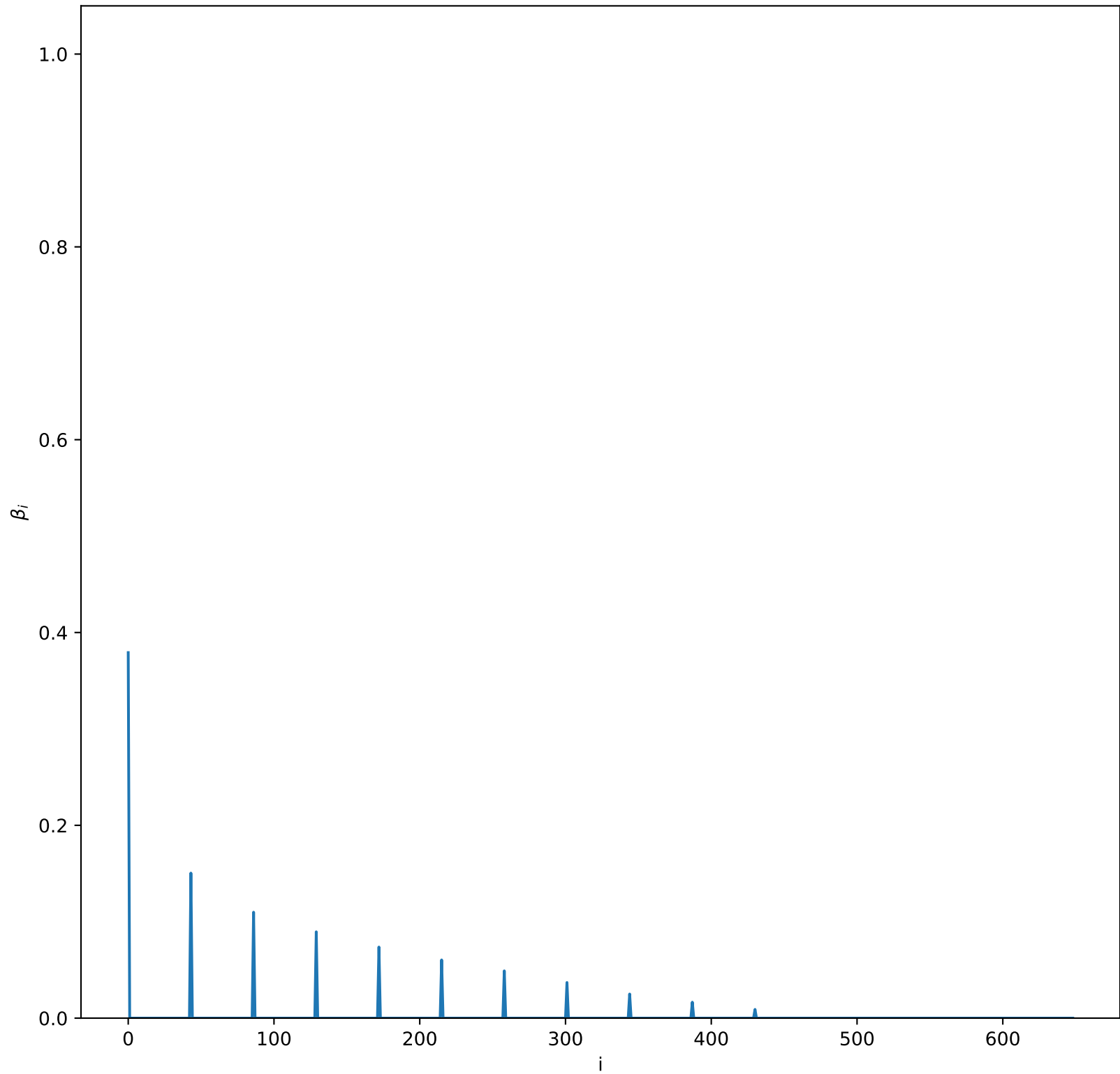
$\mu = 0.68$



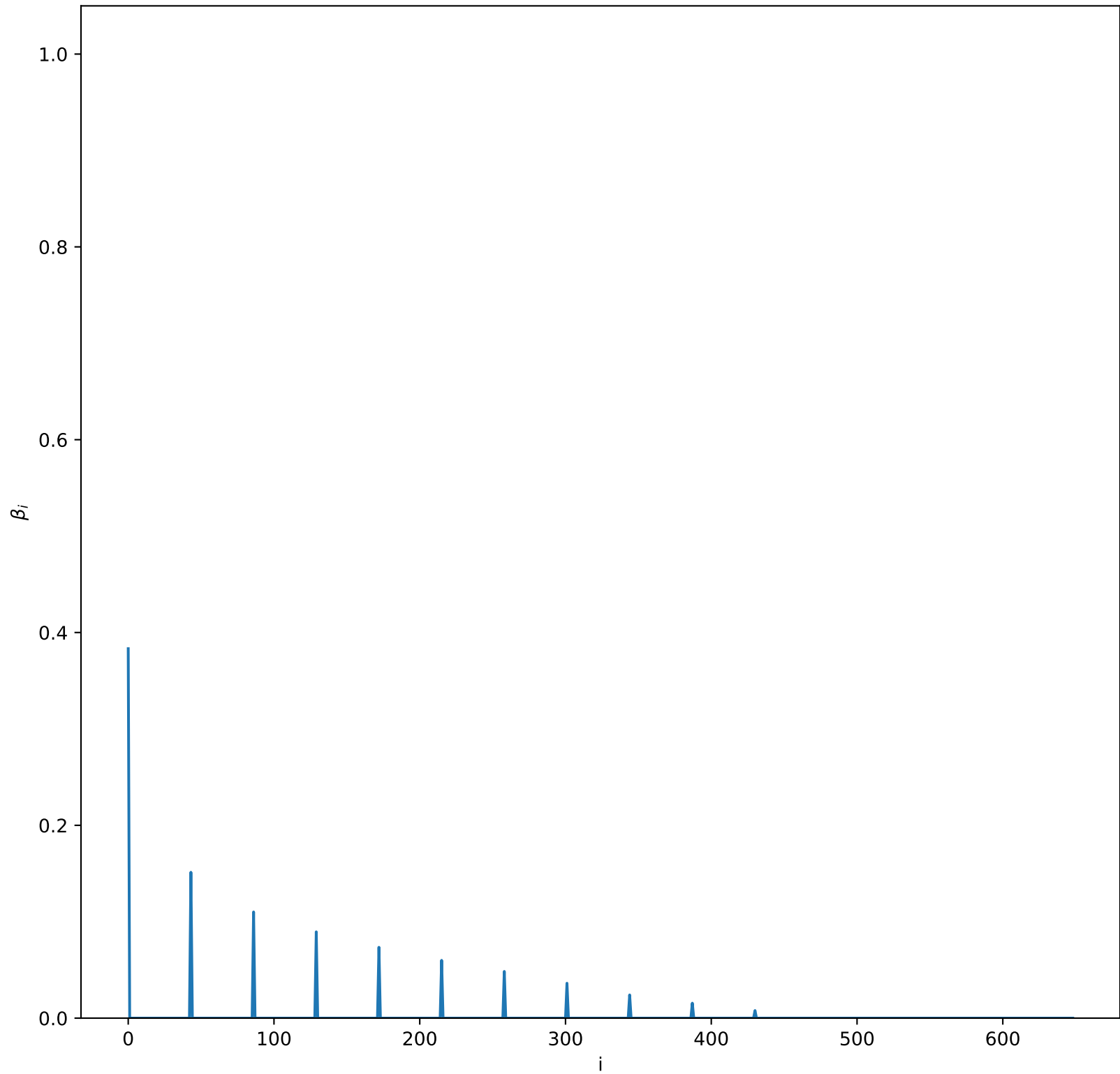
$\mu = 0.69$



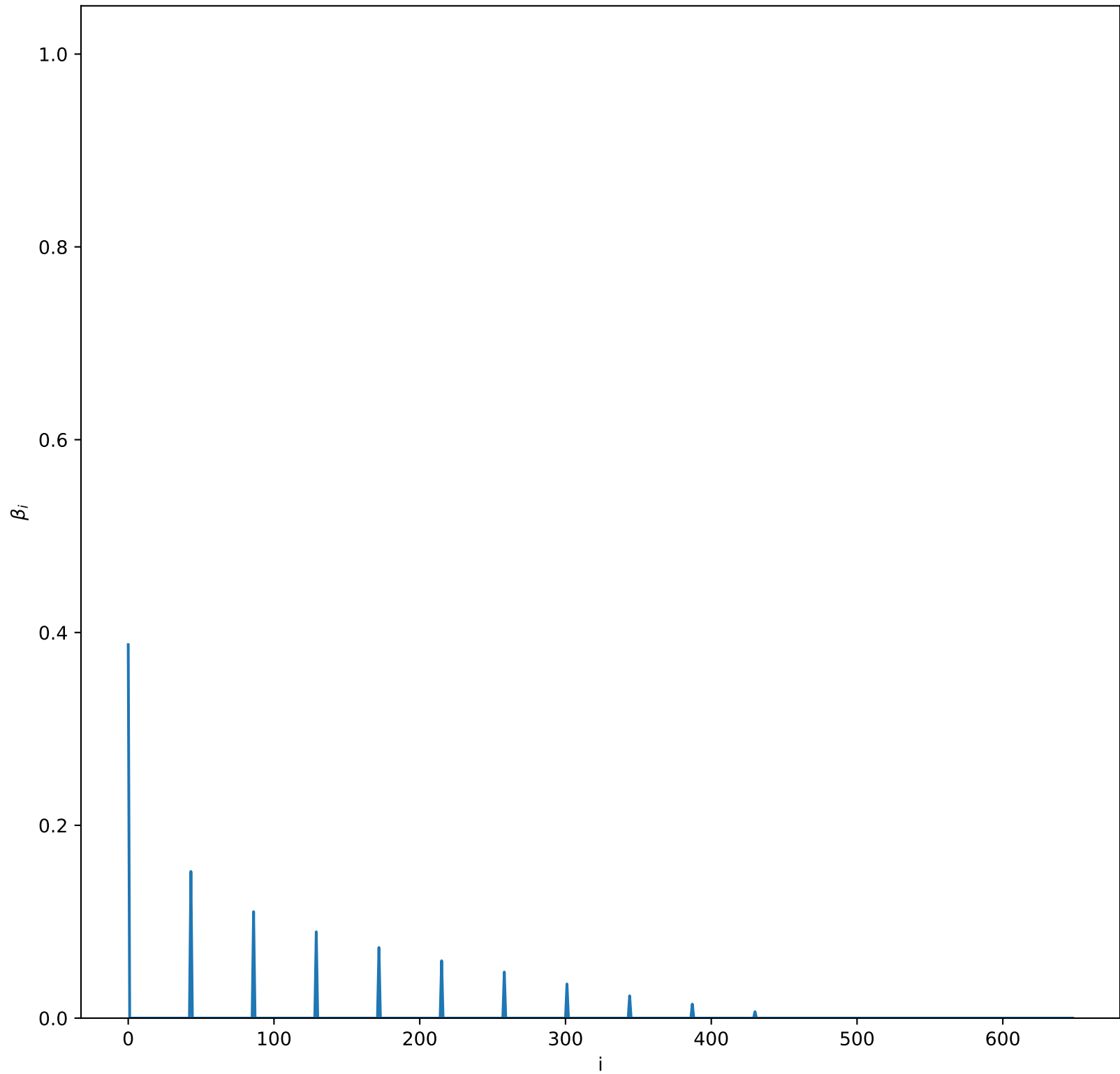
$\mu = 0.70$



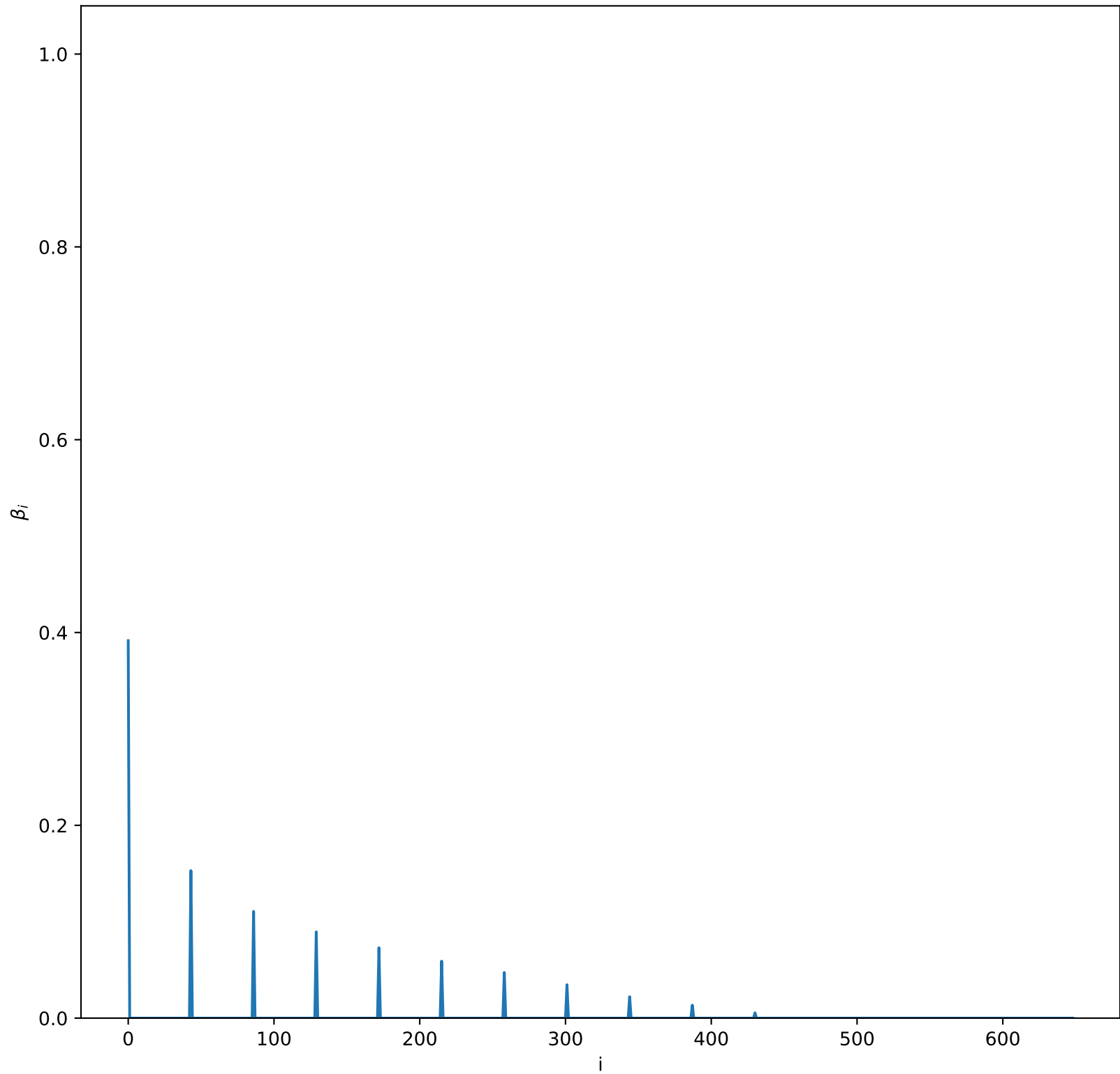
$\mu = 0.71$



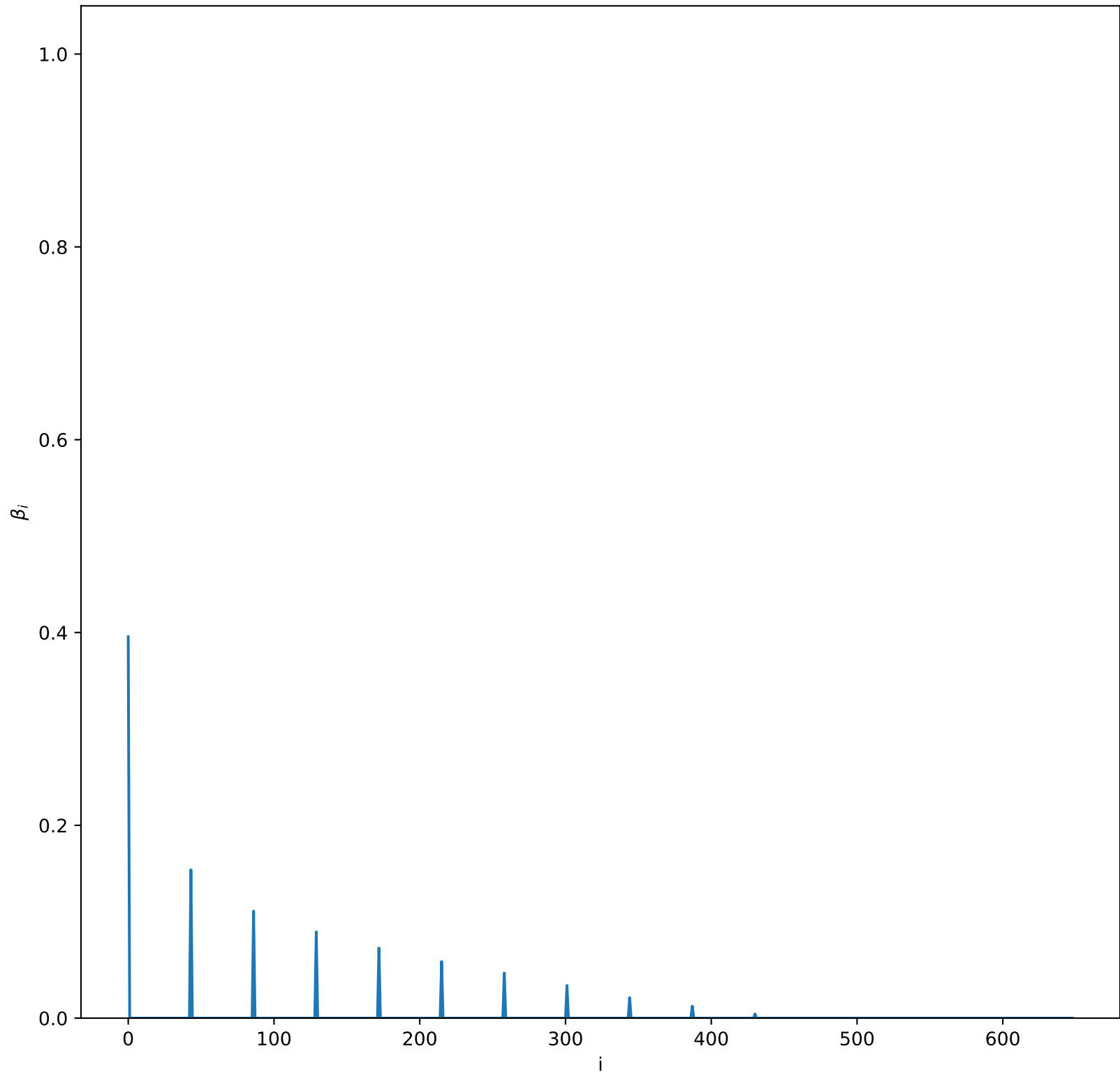
$\mu = 0.72$



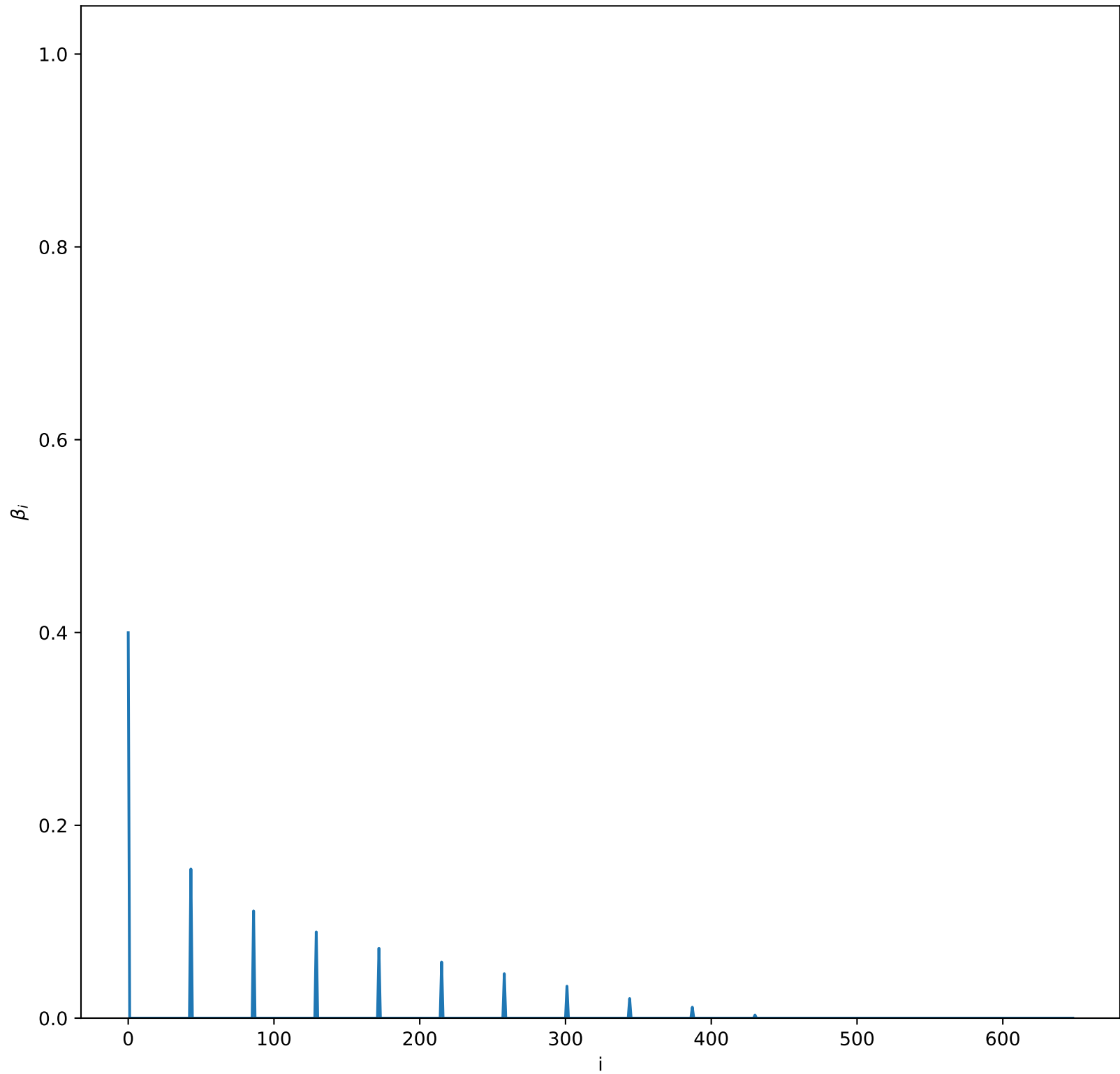
$\mu = 0.73$



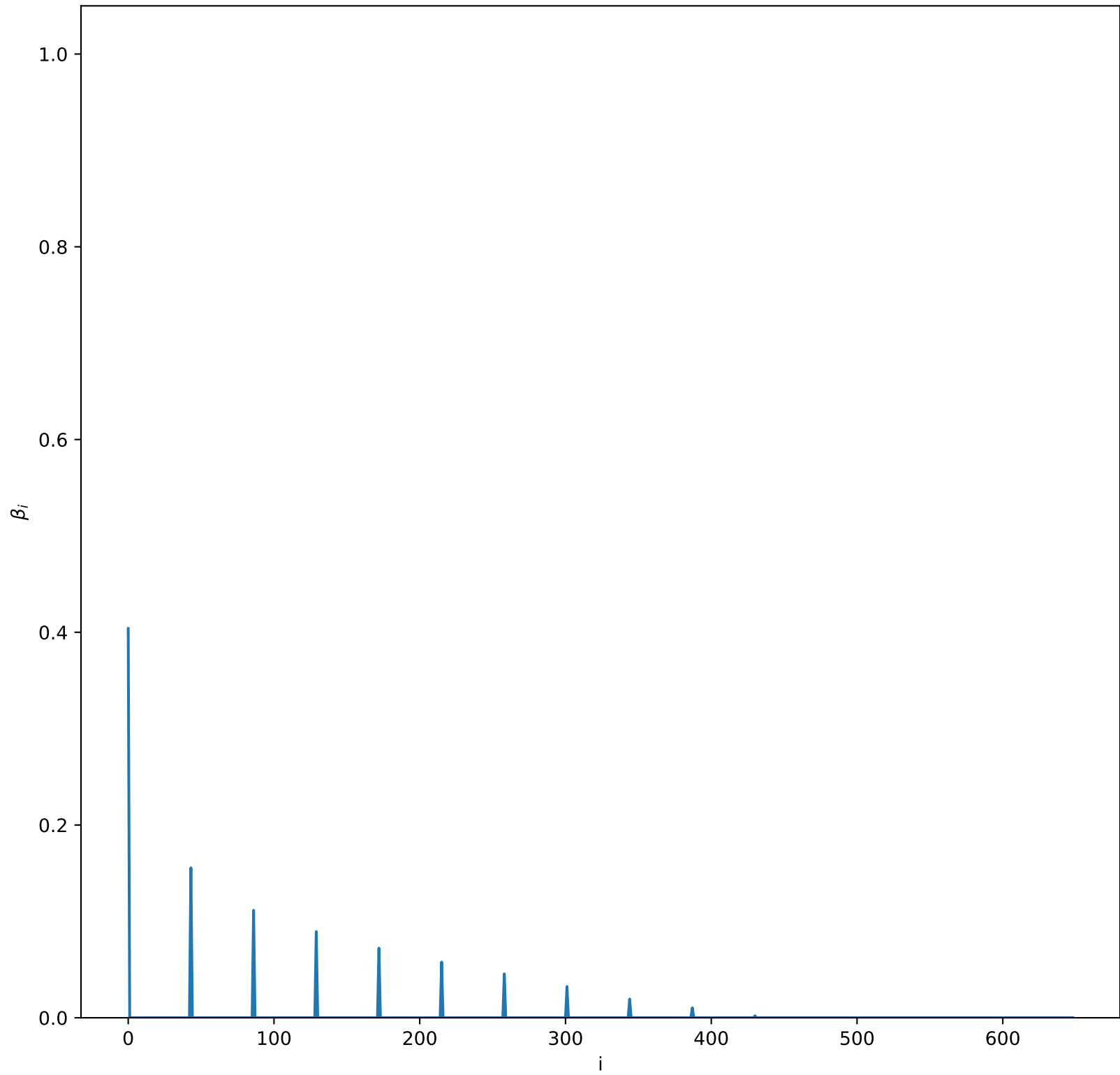
$\mu = 0.74$



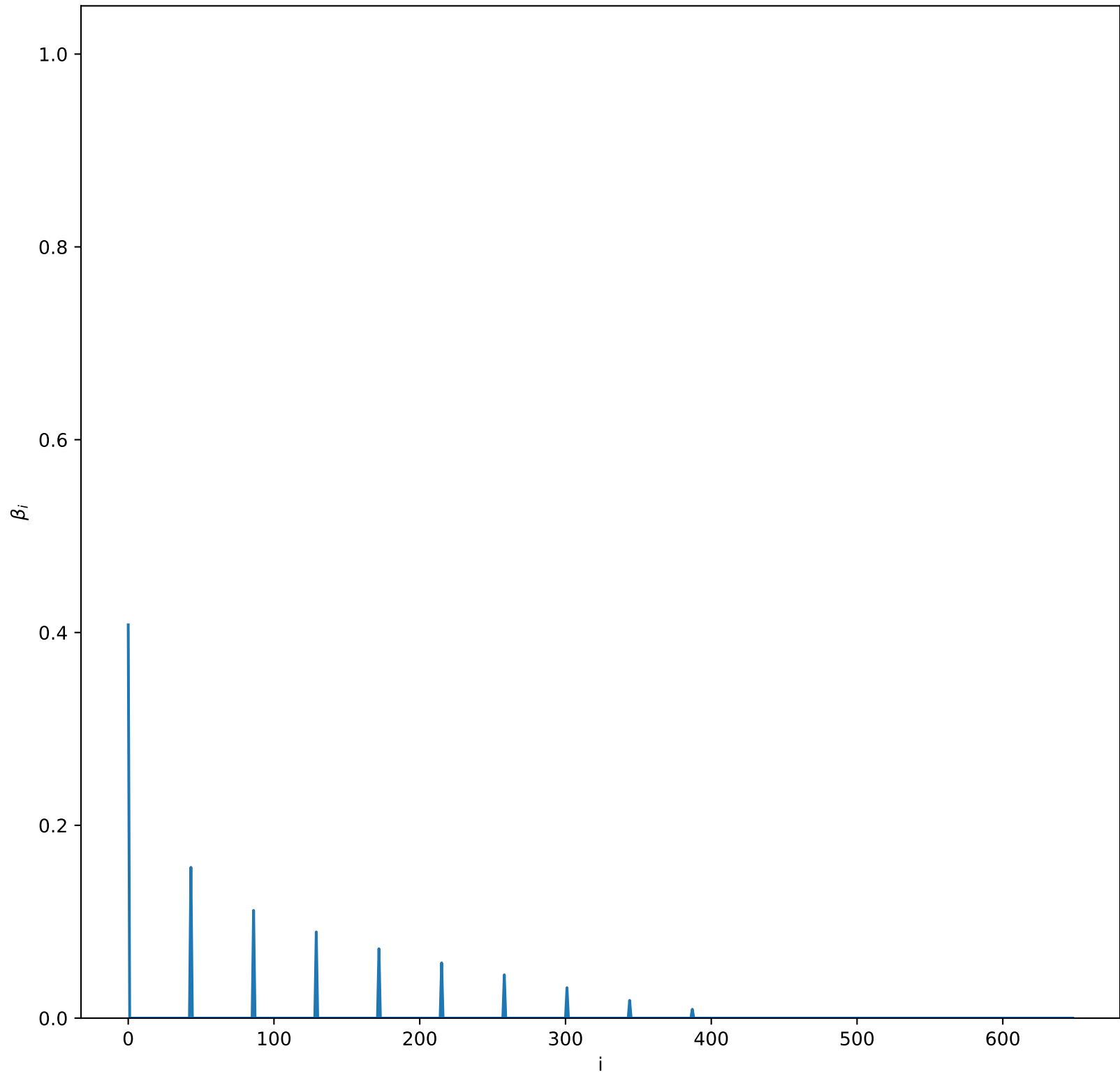
$\mu = 0.75$



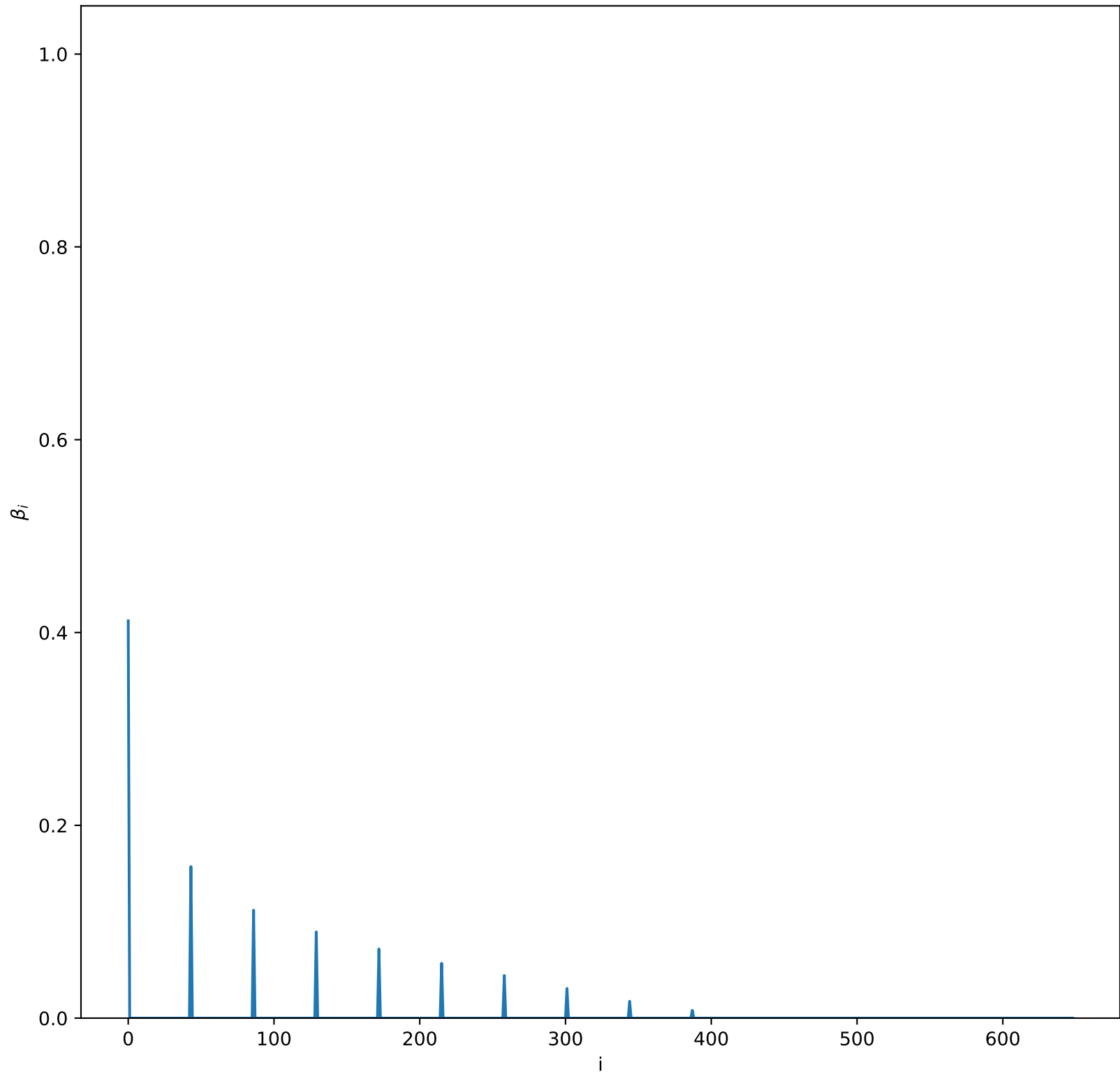
$\mu = 0.76$



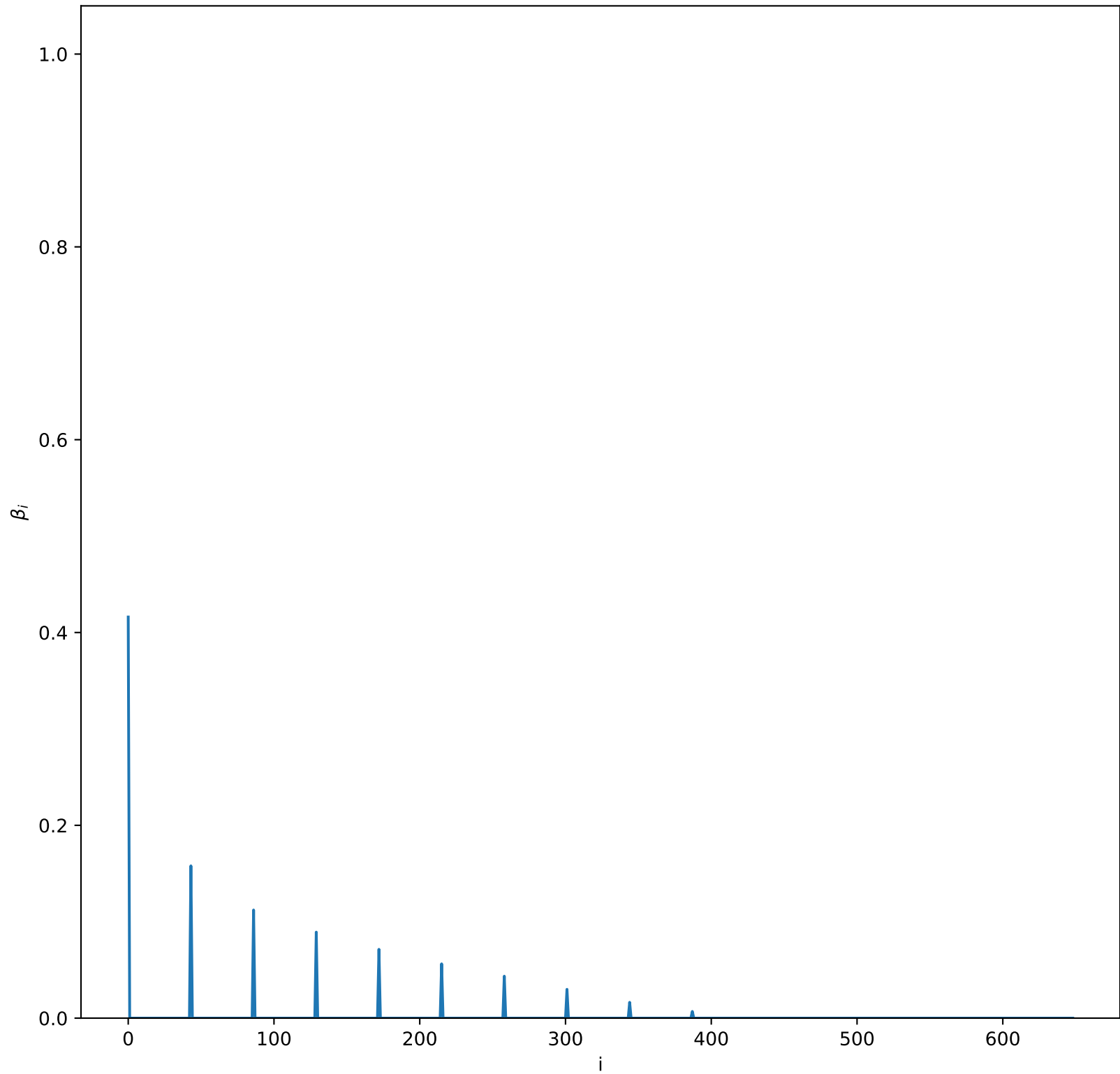
$\mu = 0.77$



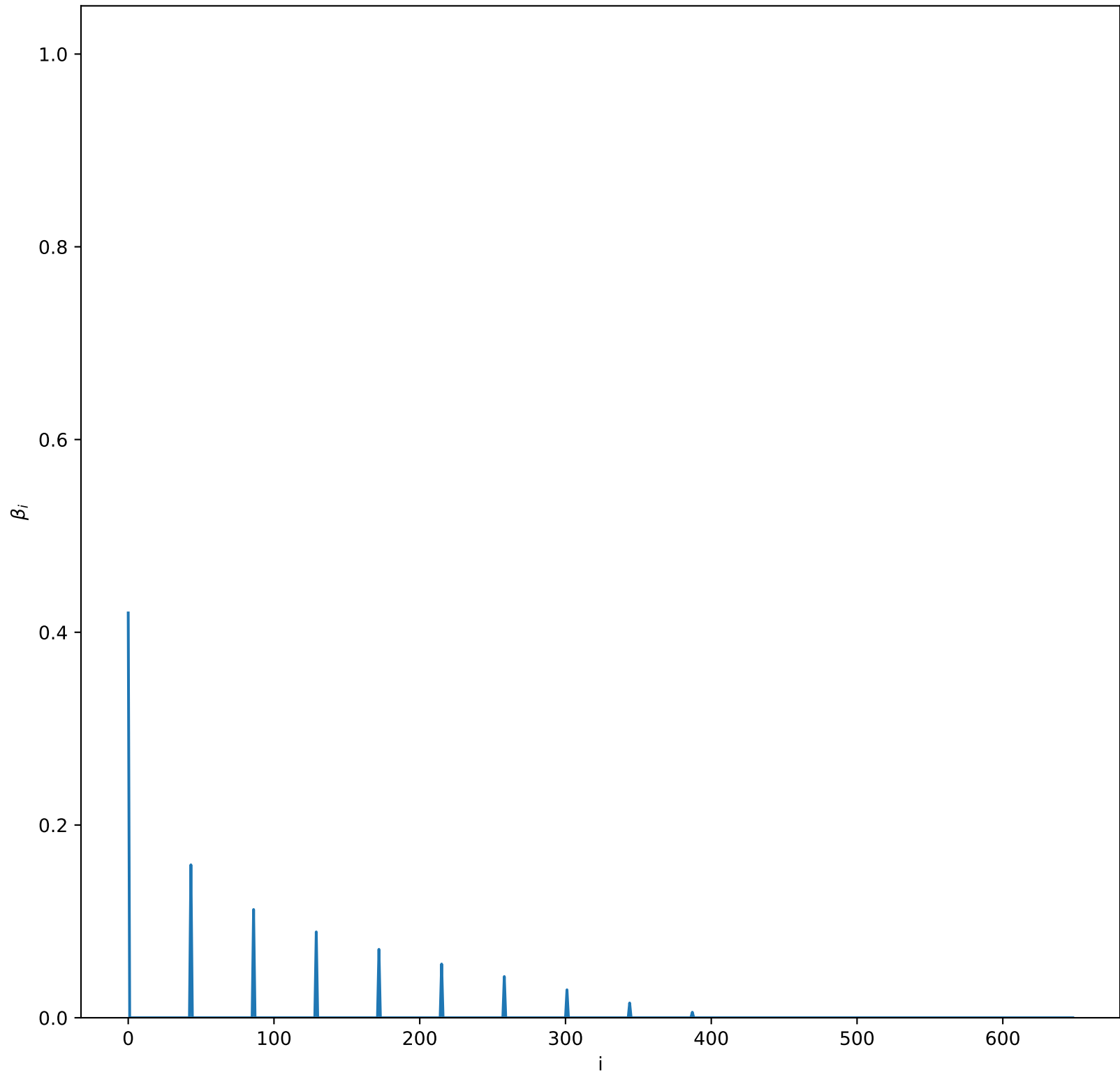
$\mu = 0.78$



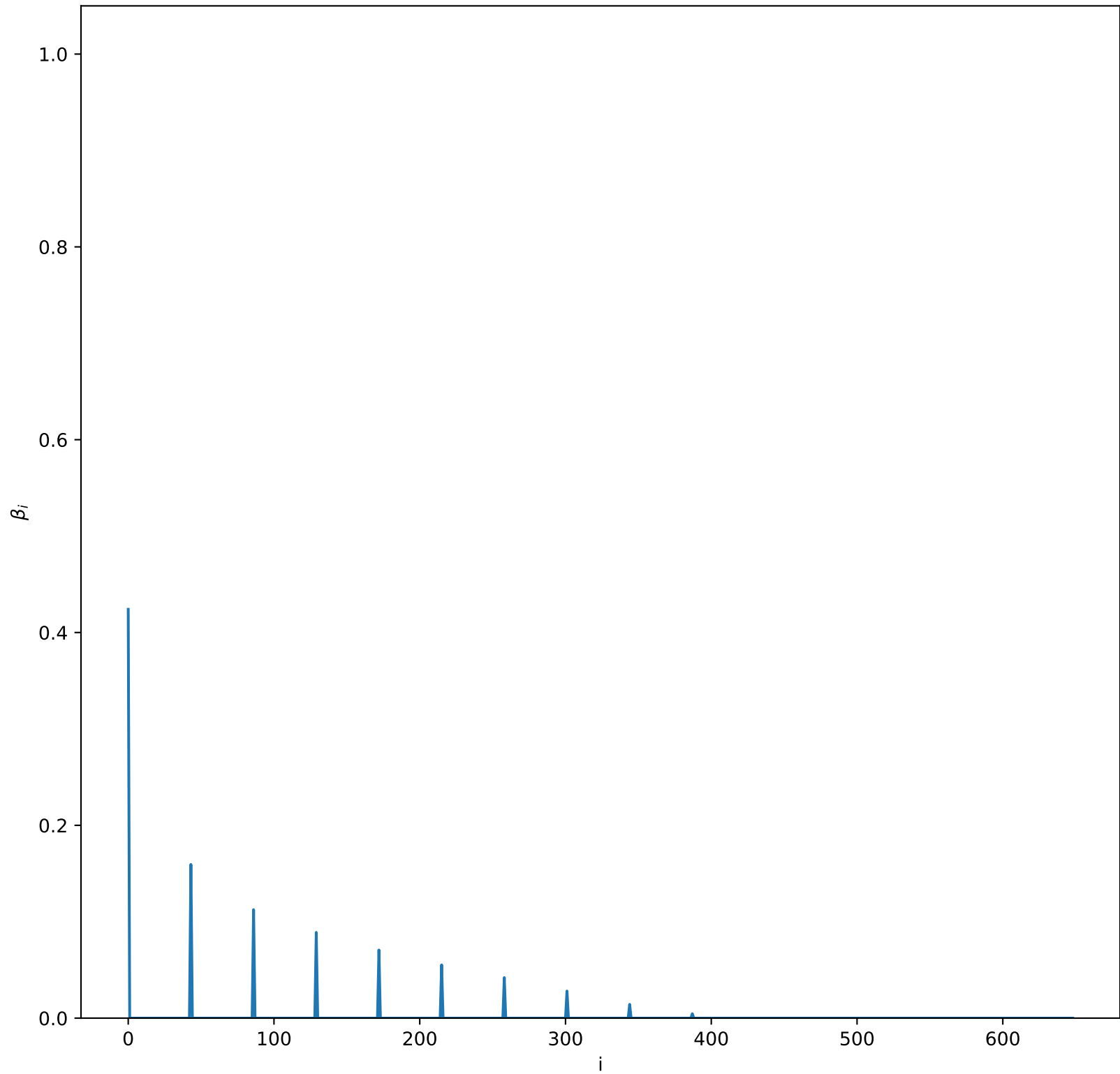
$\mu = 0.79$



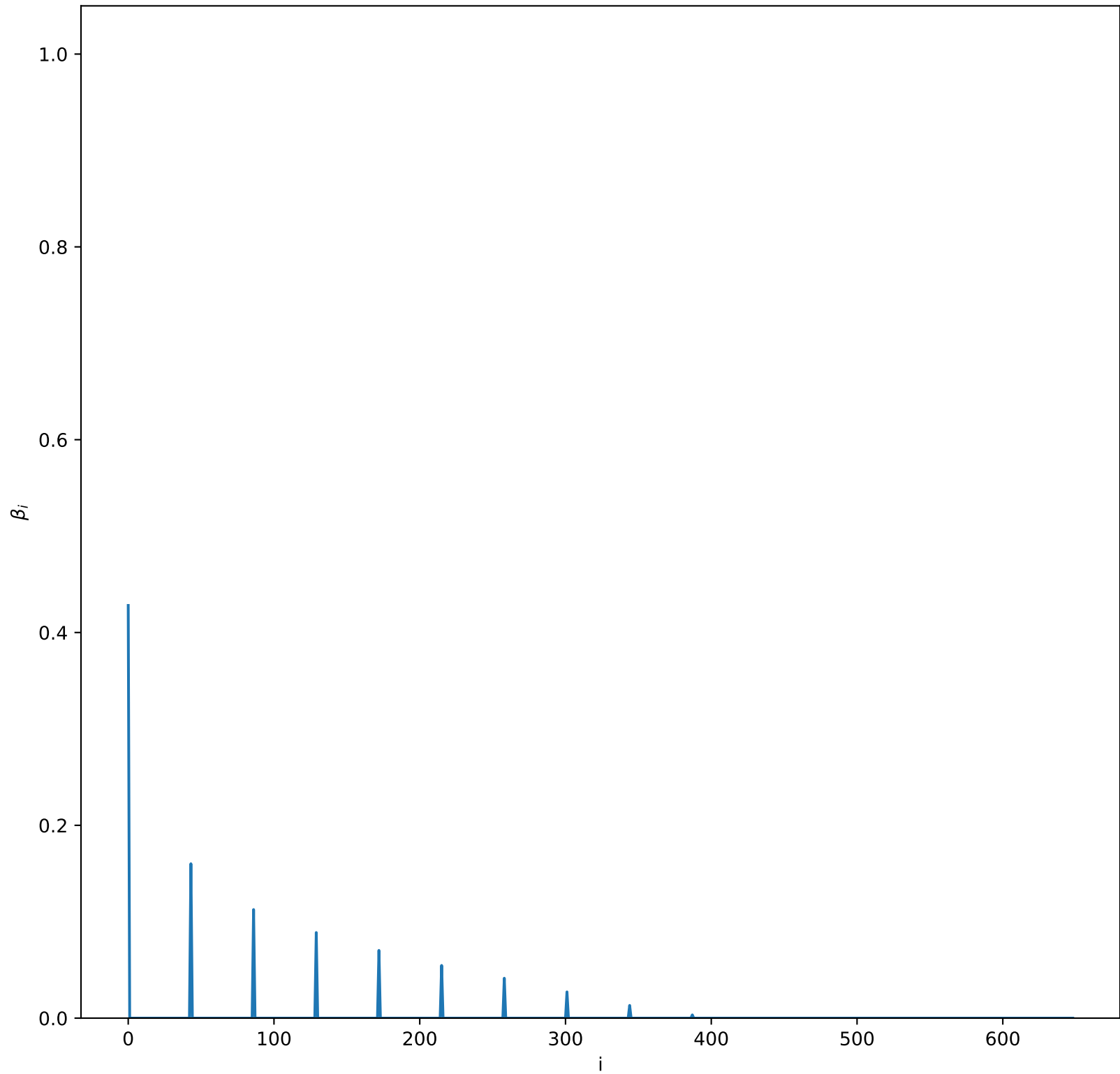
$\mu = 0.80$



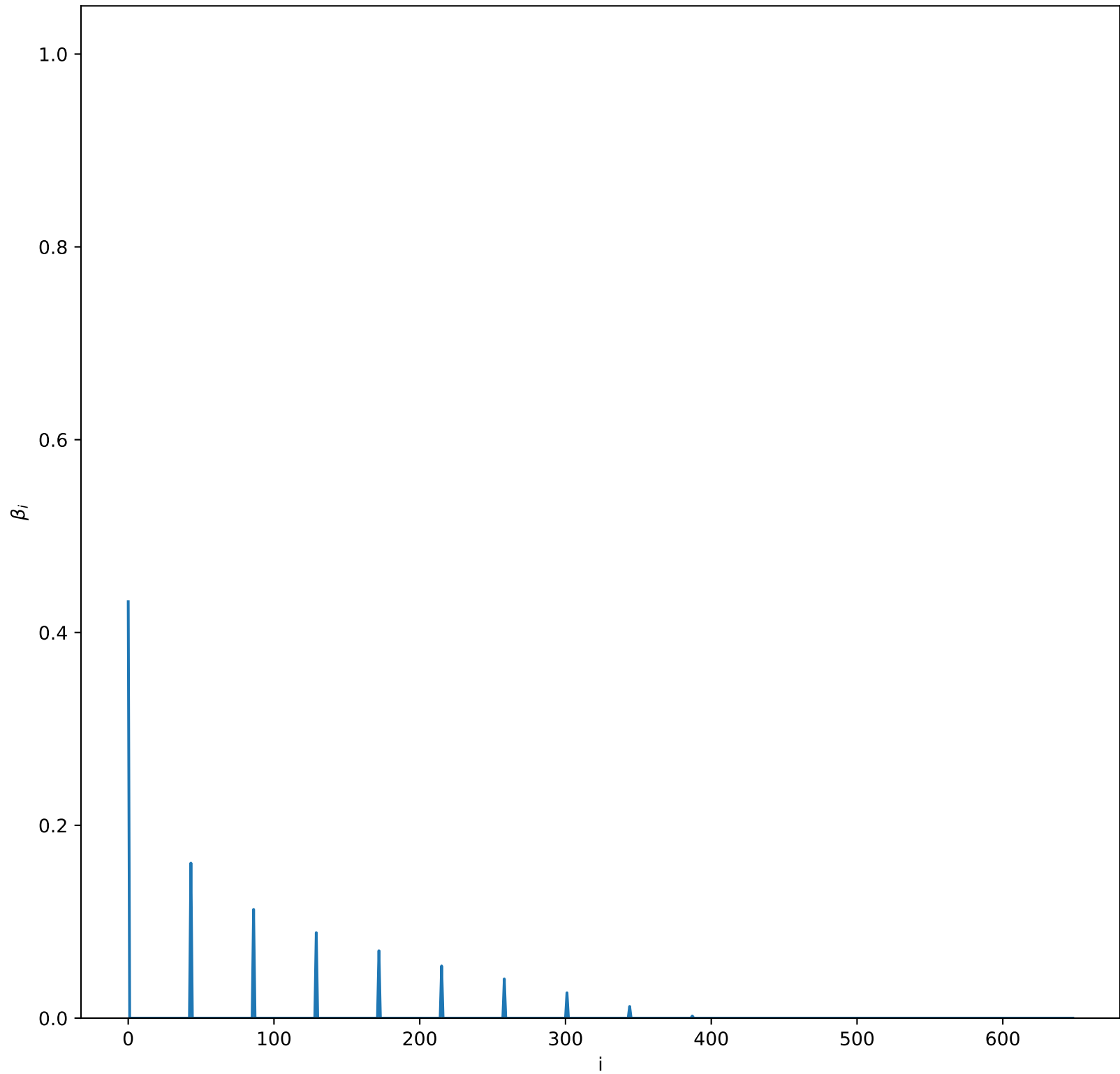
$\mu = 0.81$



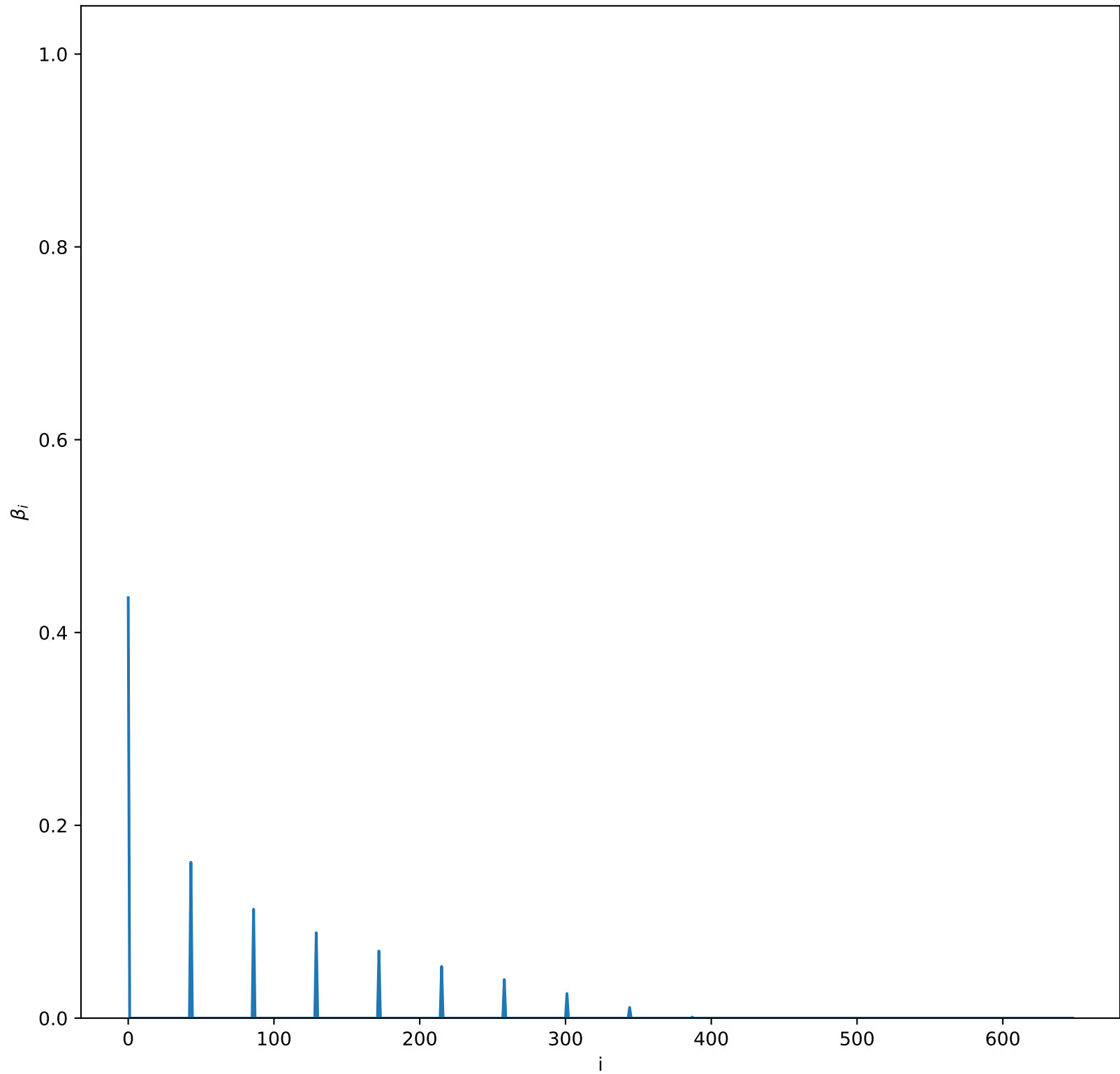
$\mu = 0.82$



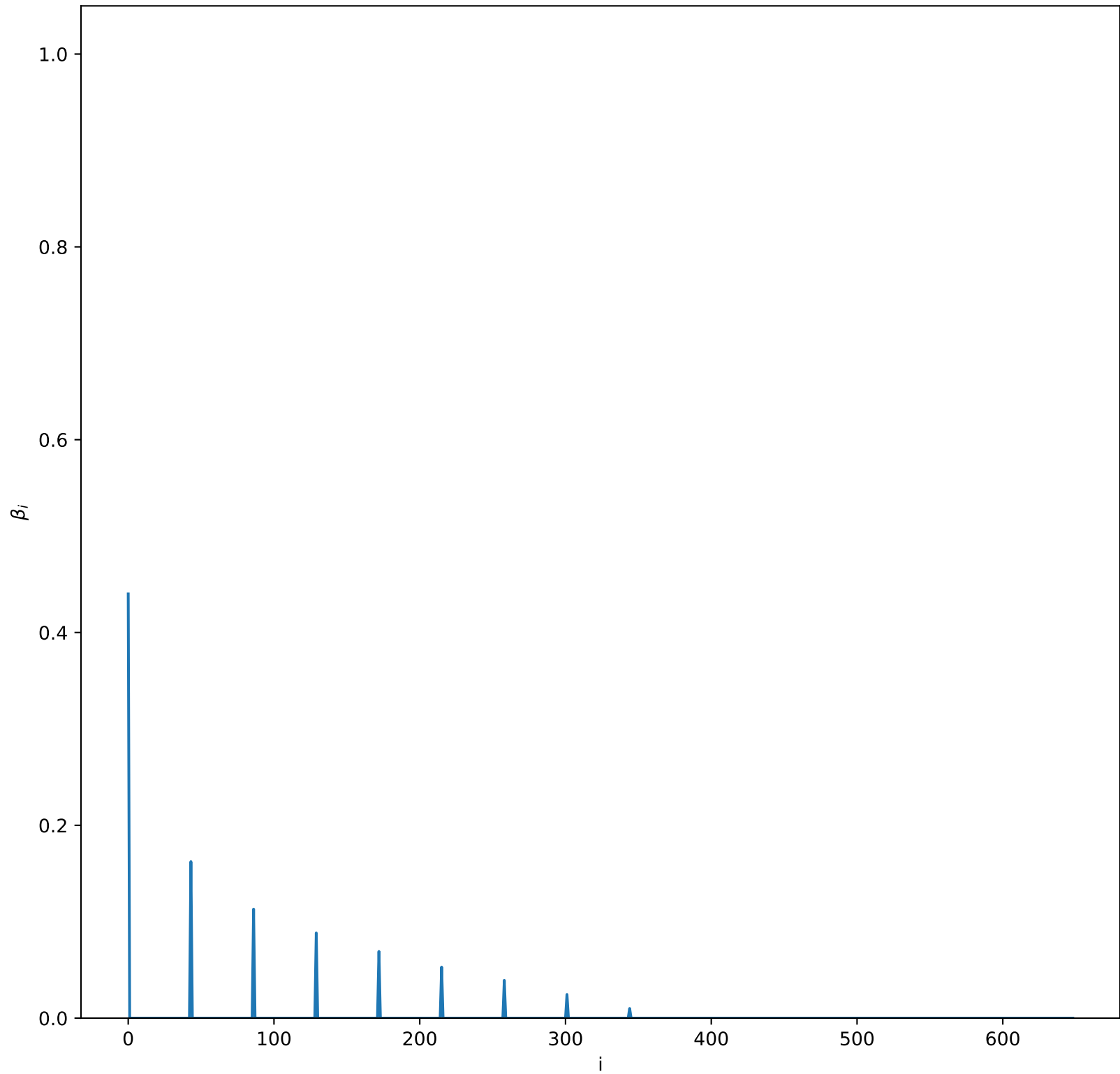
$\mu = 0.83$



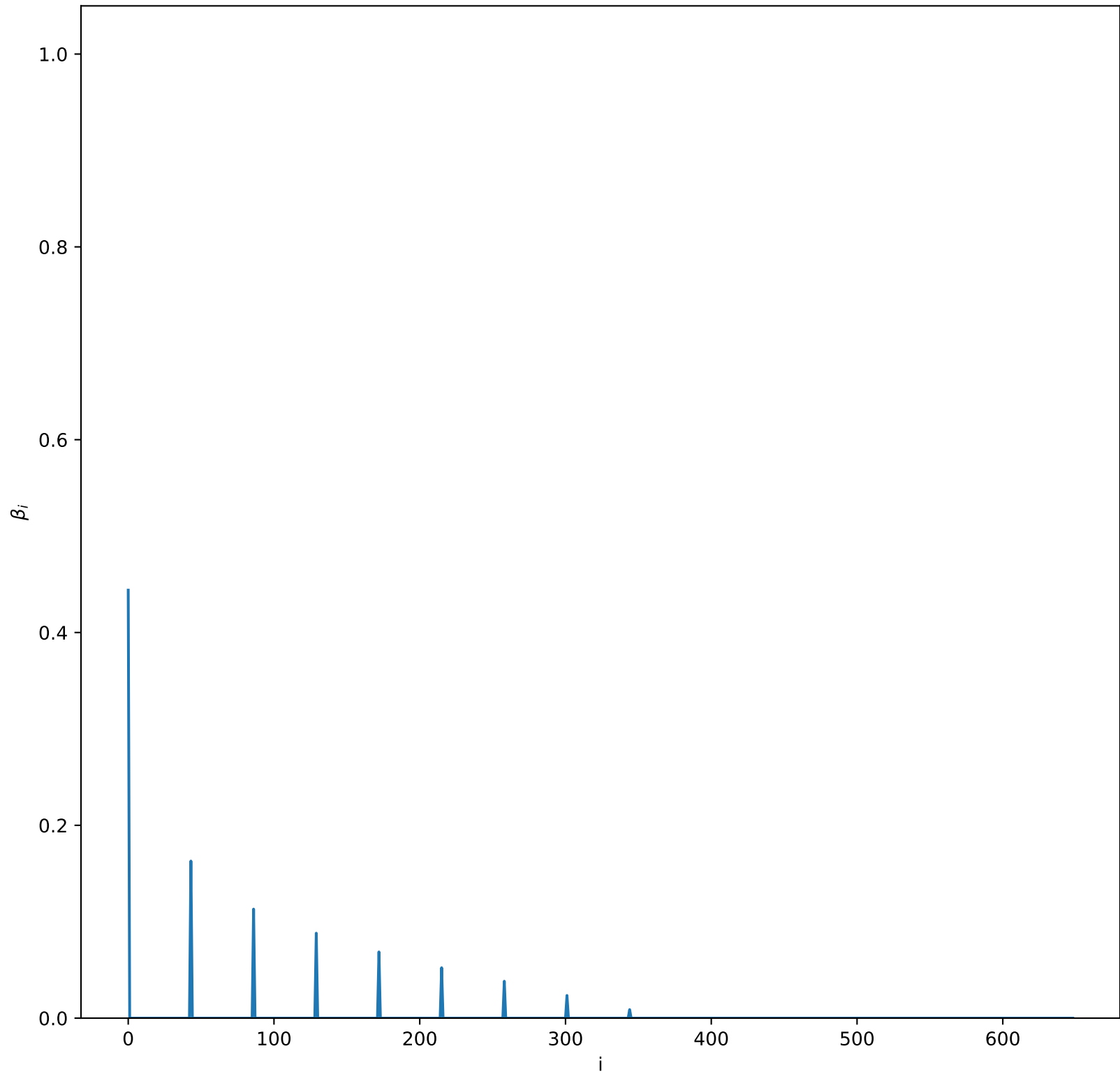
$\mu = 0.84$



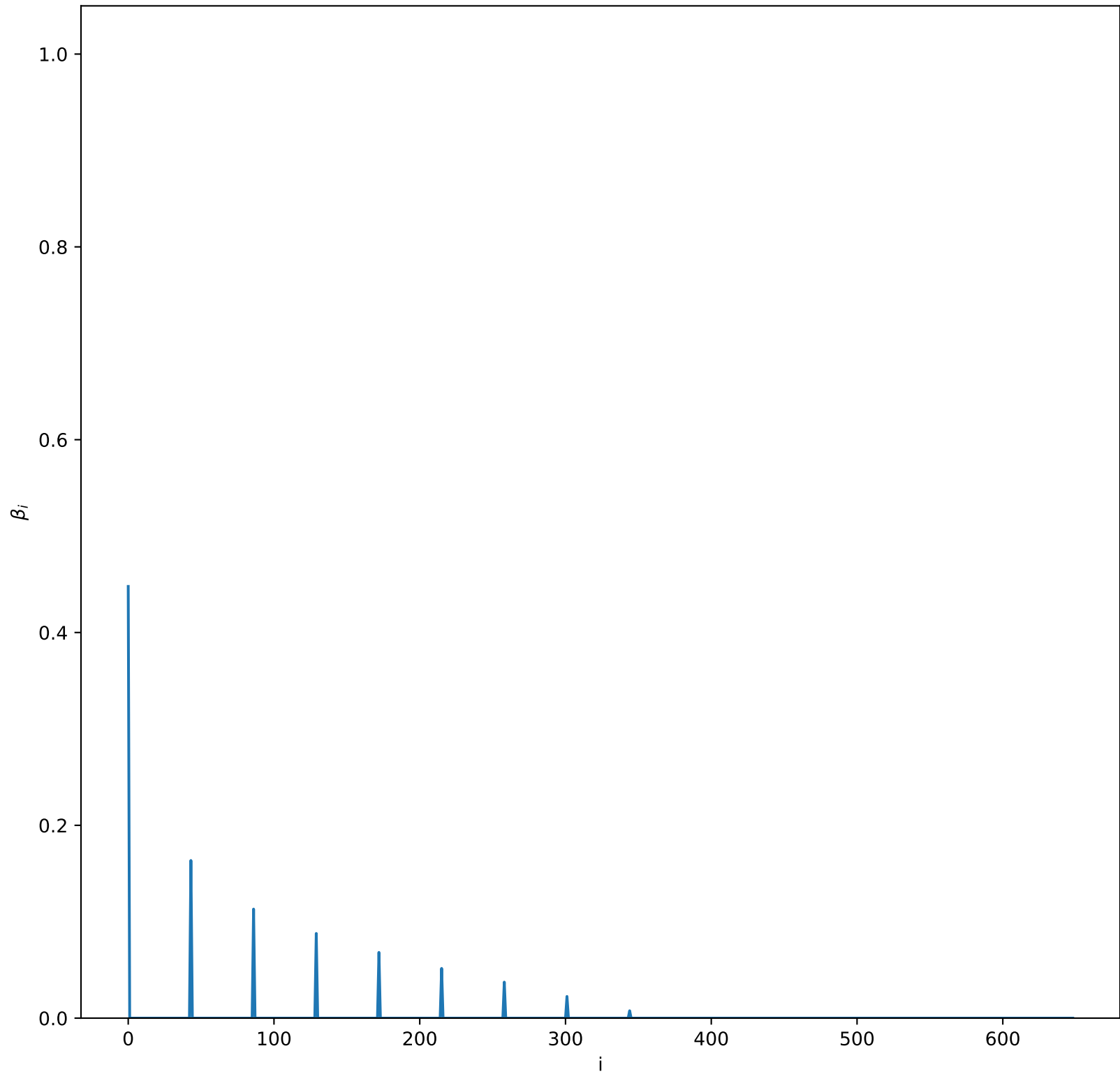
$\mu = 0.85$



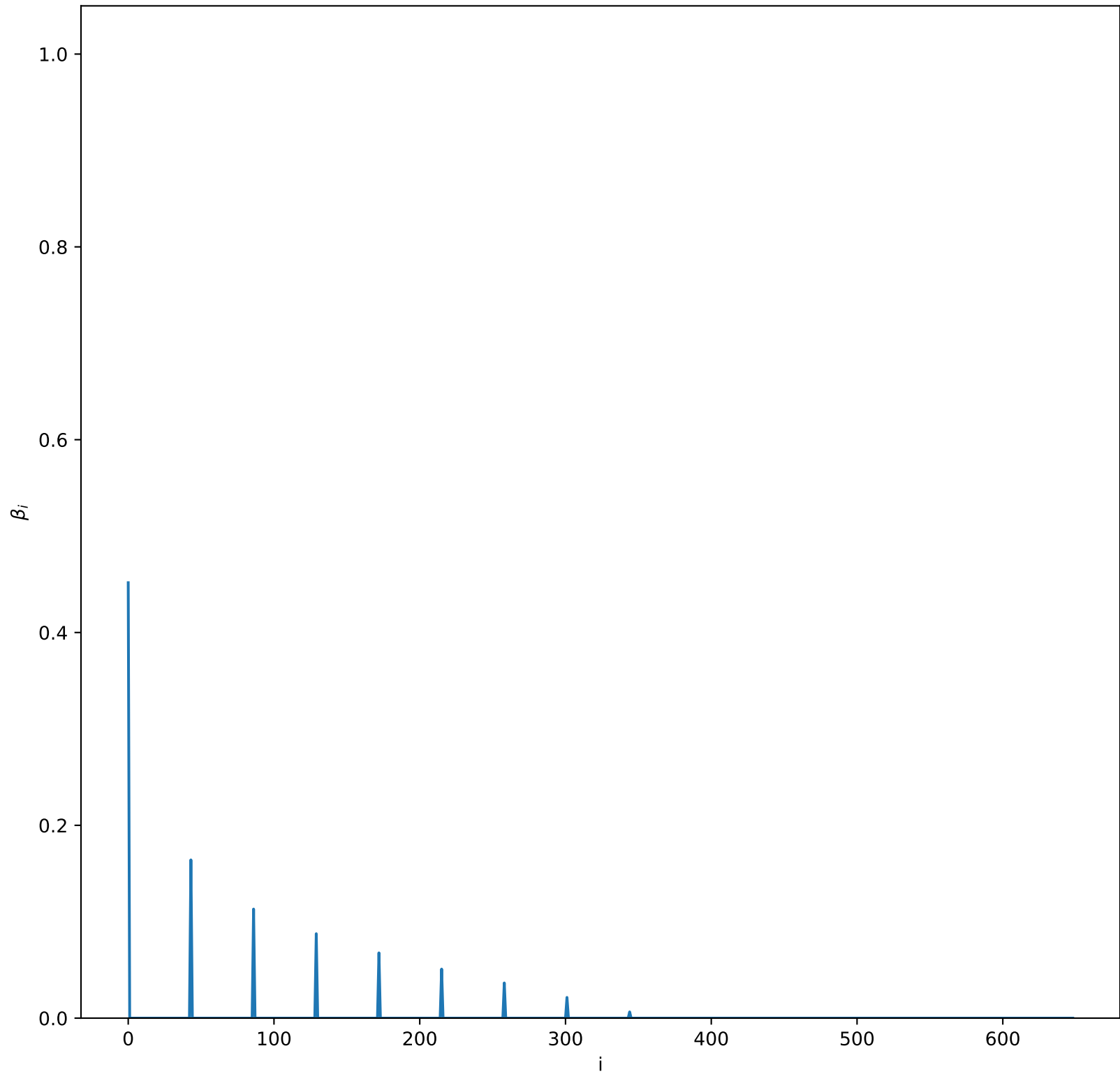
$\mu = 0.86$



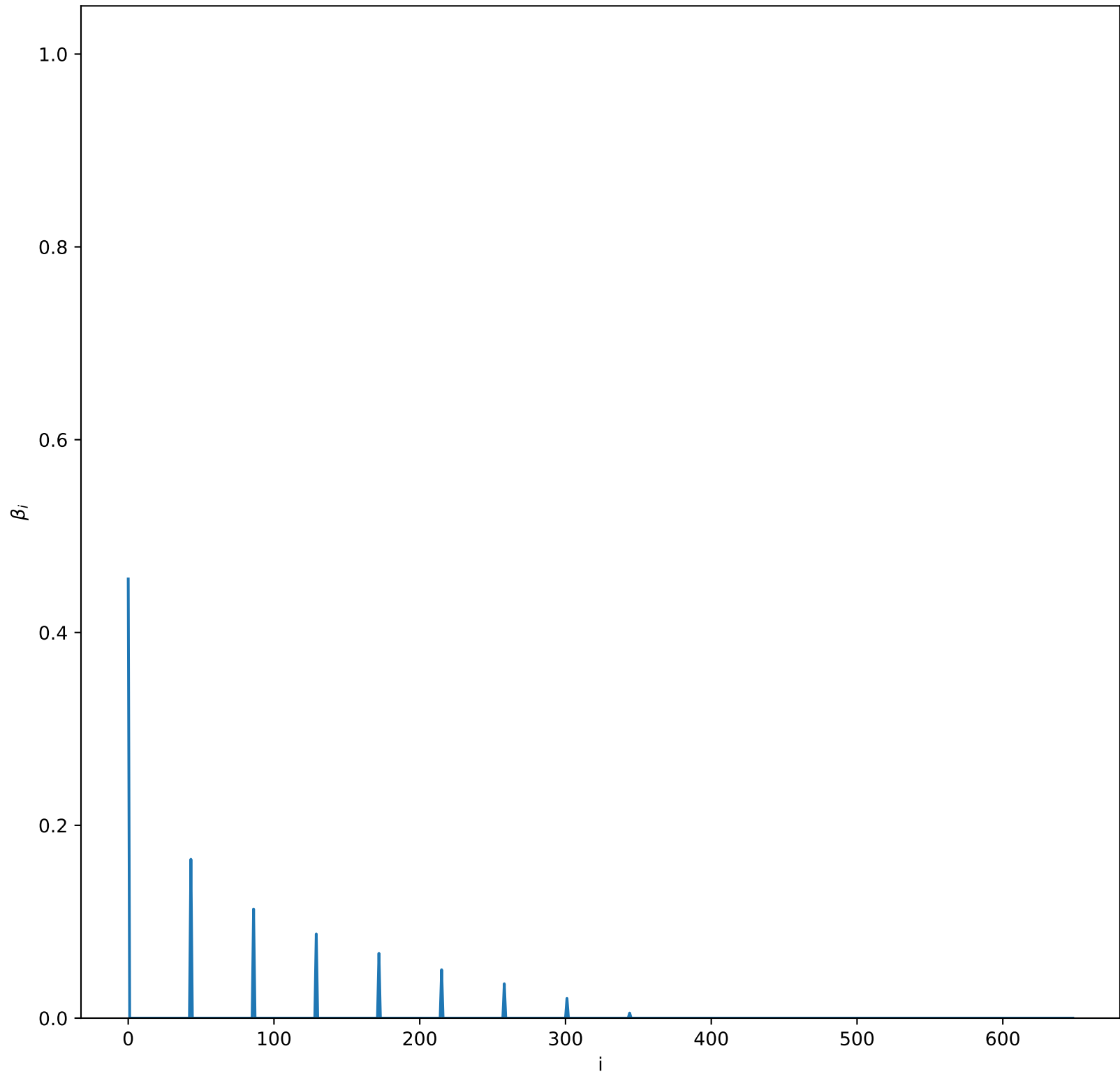
$\mu = 0.87$



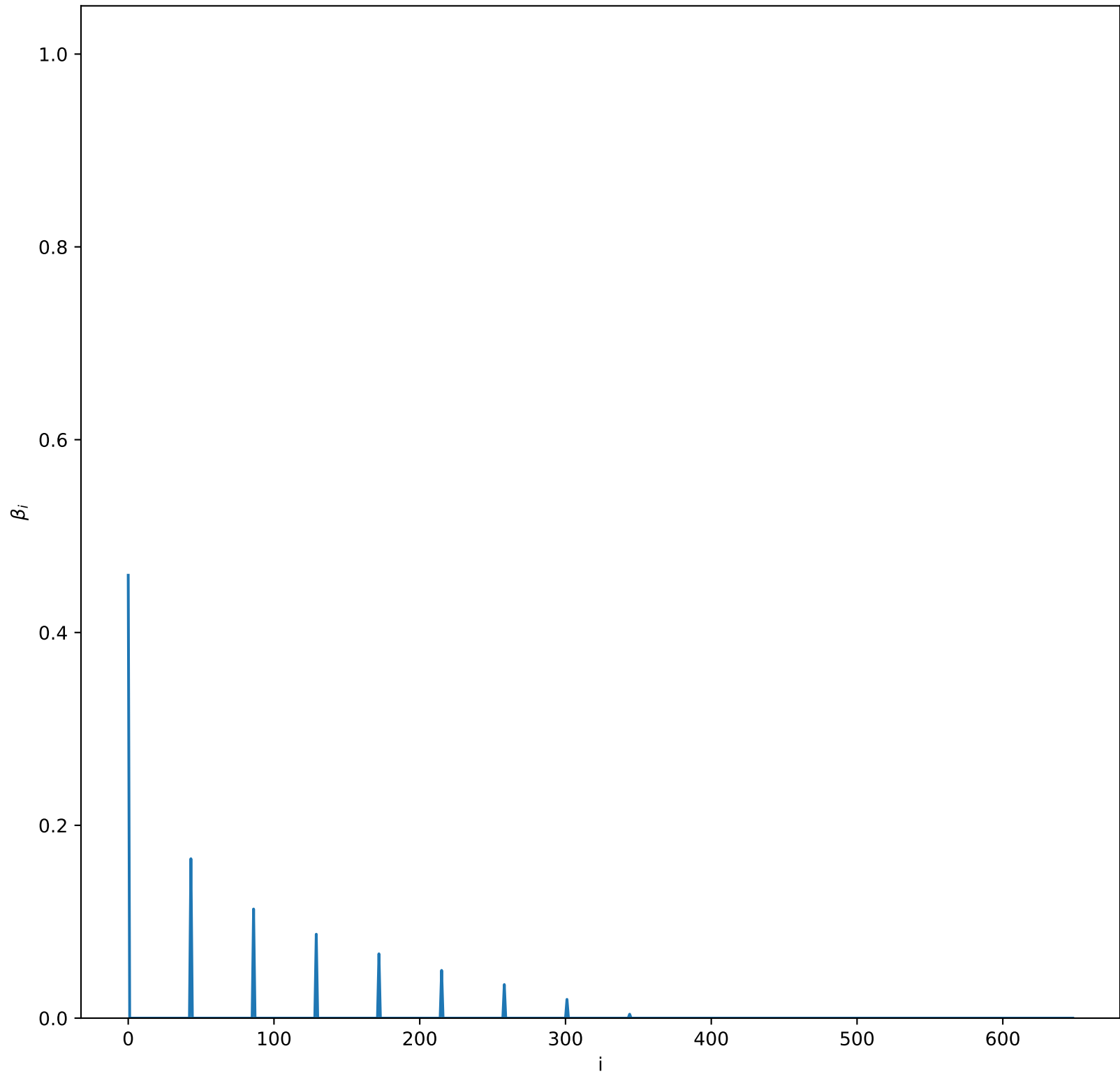
$\mu = 0.88$



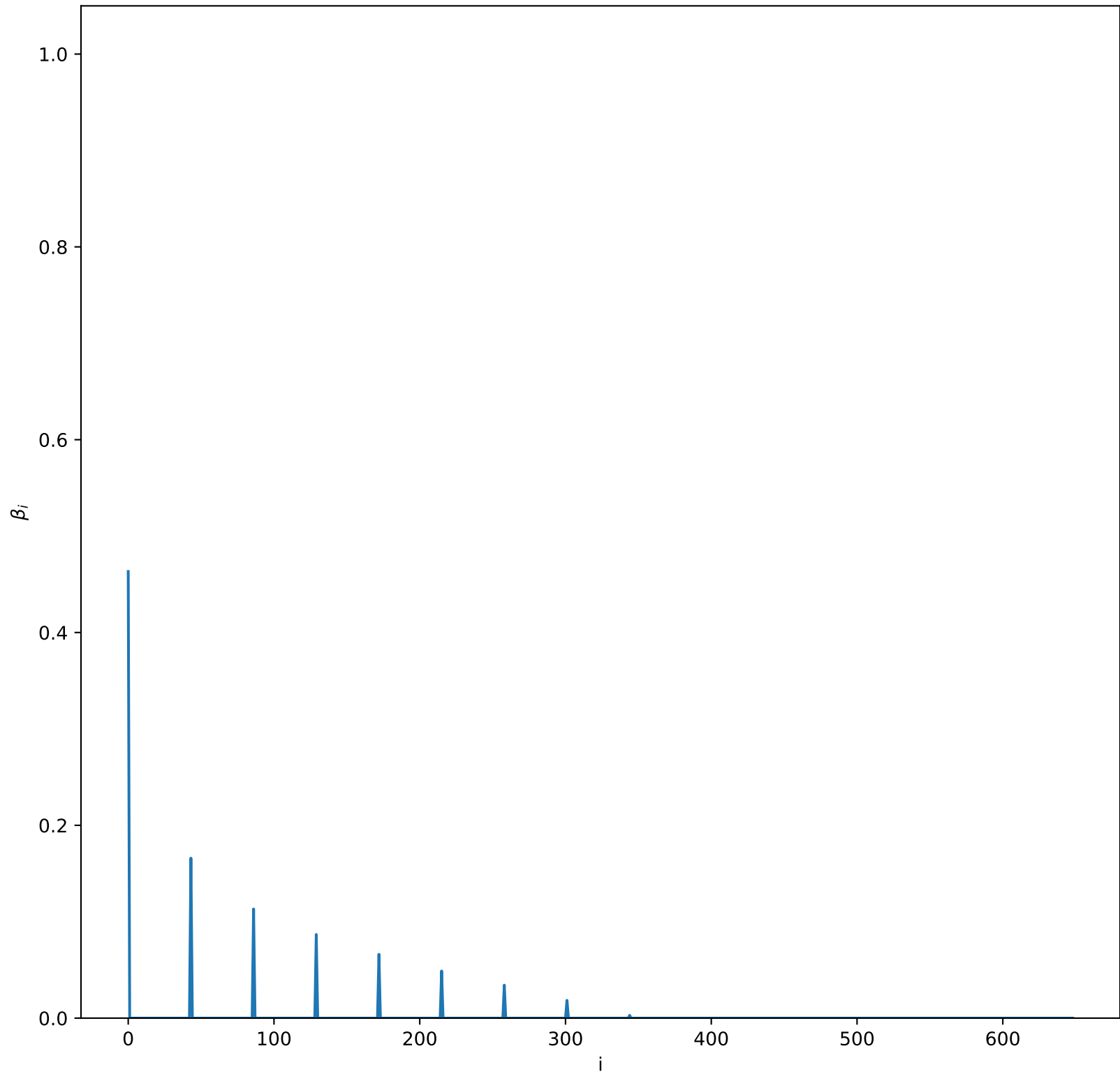
$\mu = 0.89$



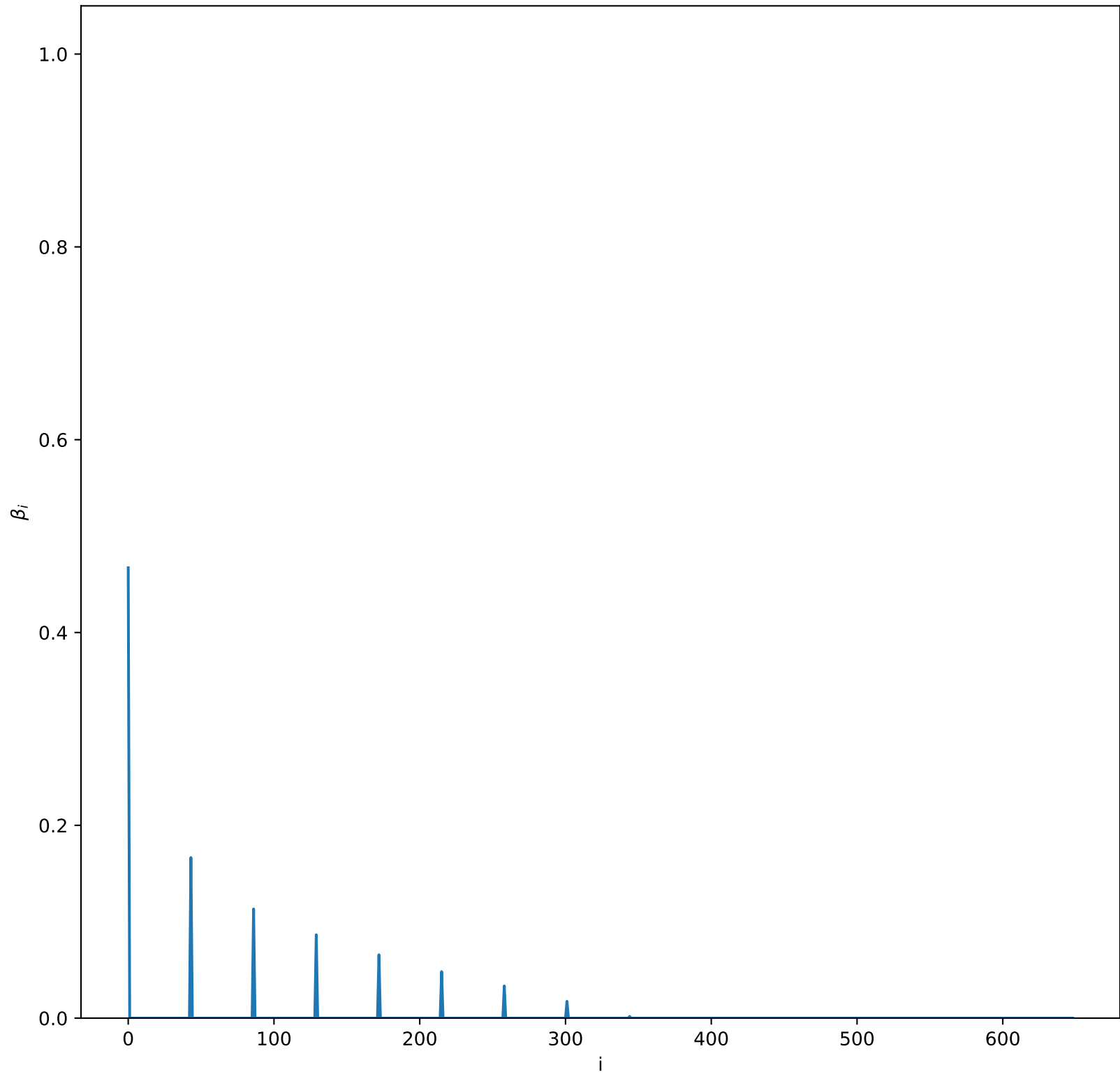
$\mu = 0.90$



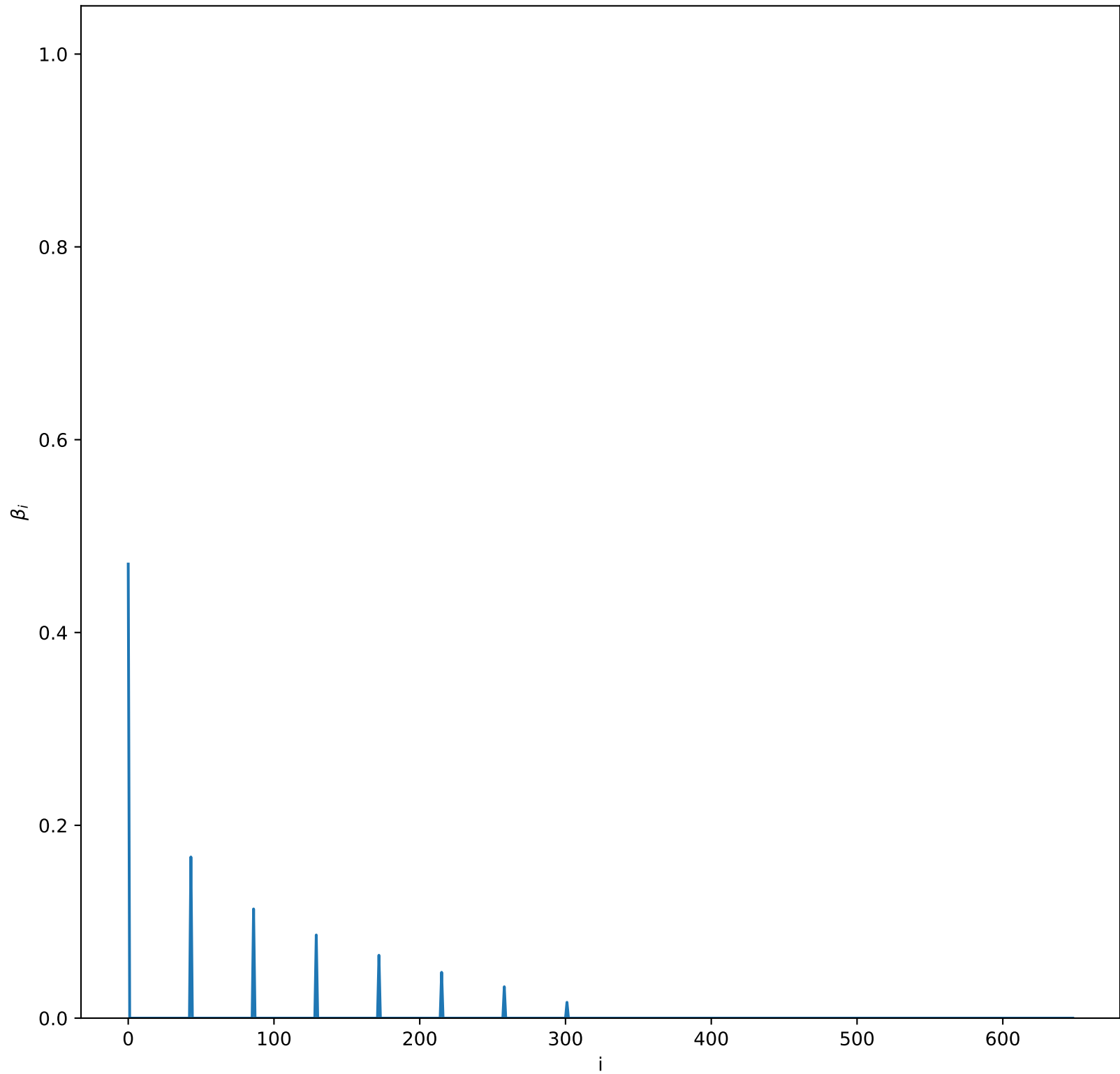
$\mu = 0.91$



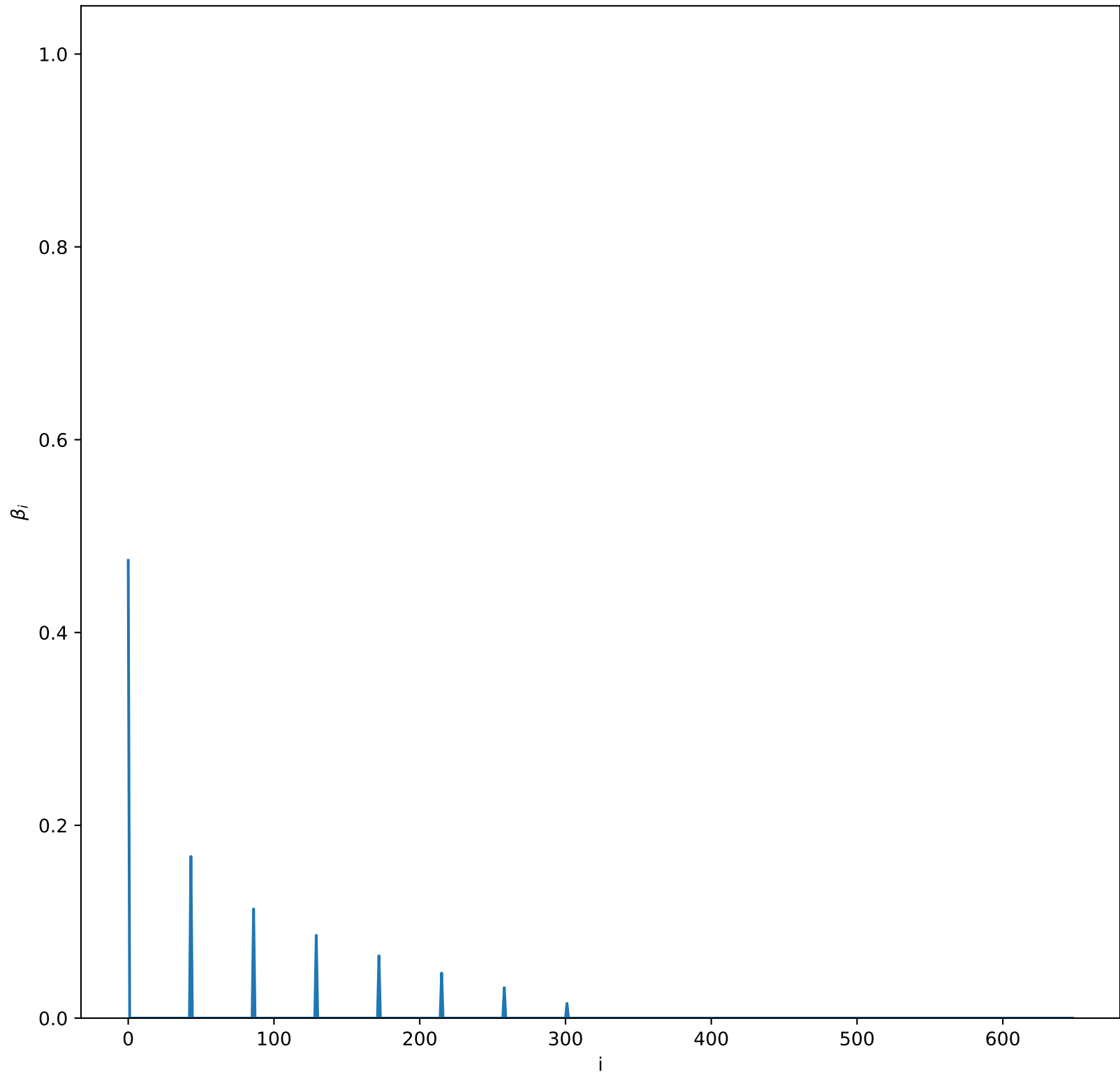
$\mu = 0.92$



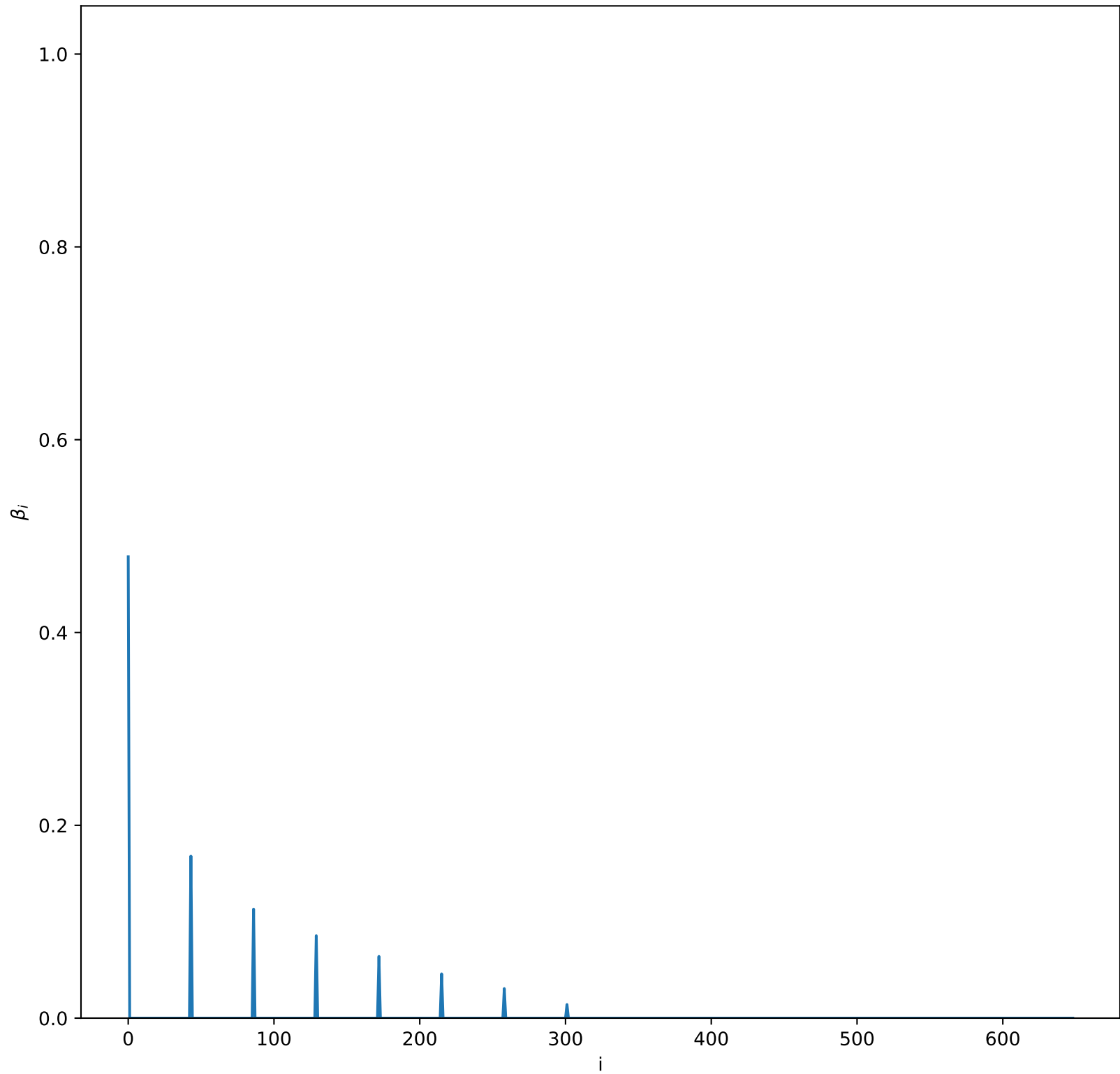
$\mu = 0.93$



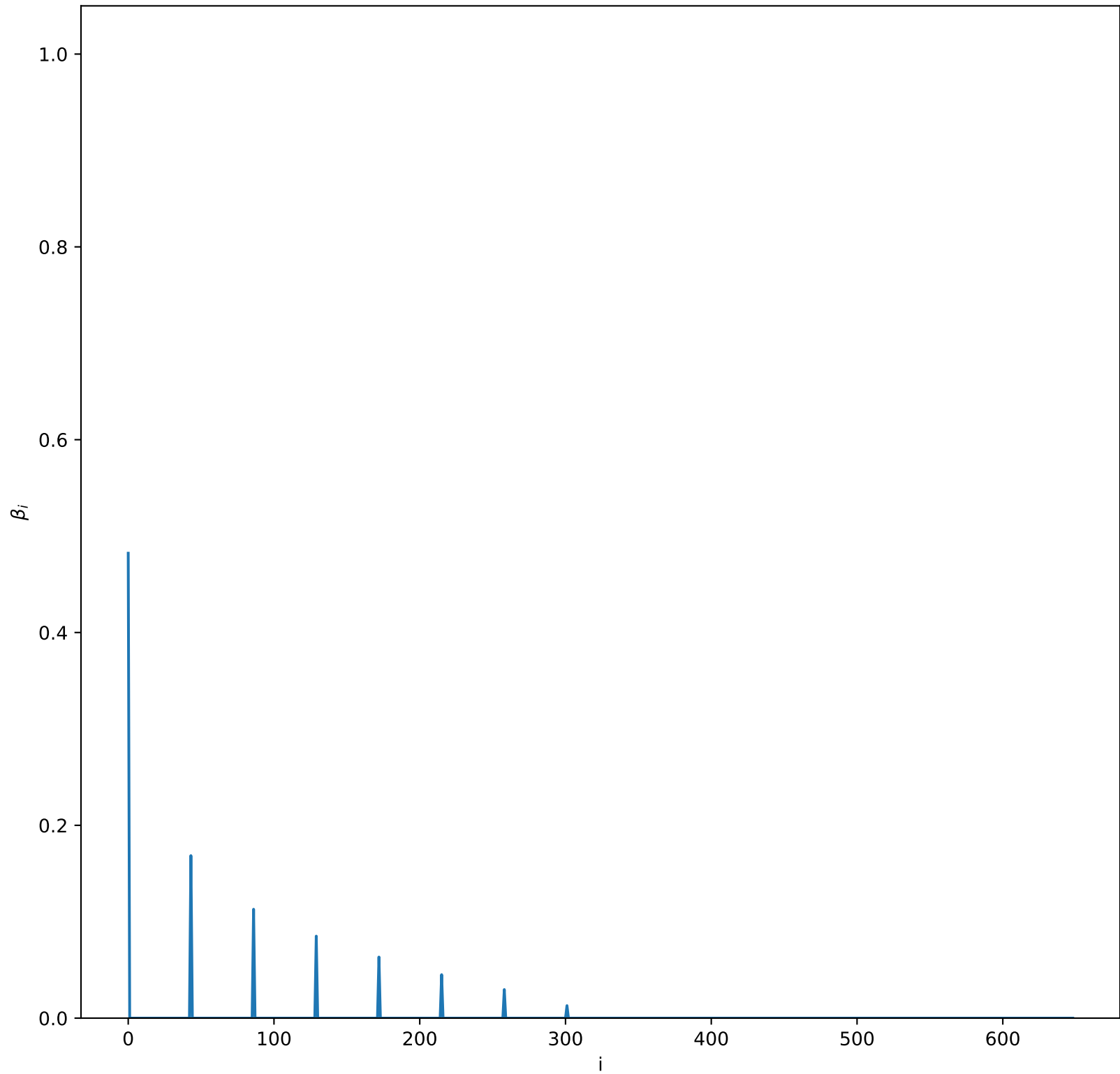
$\mu = 0.94$



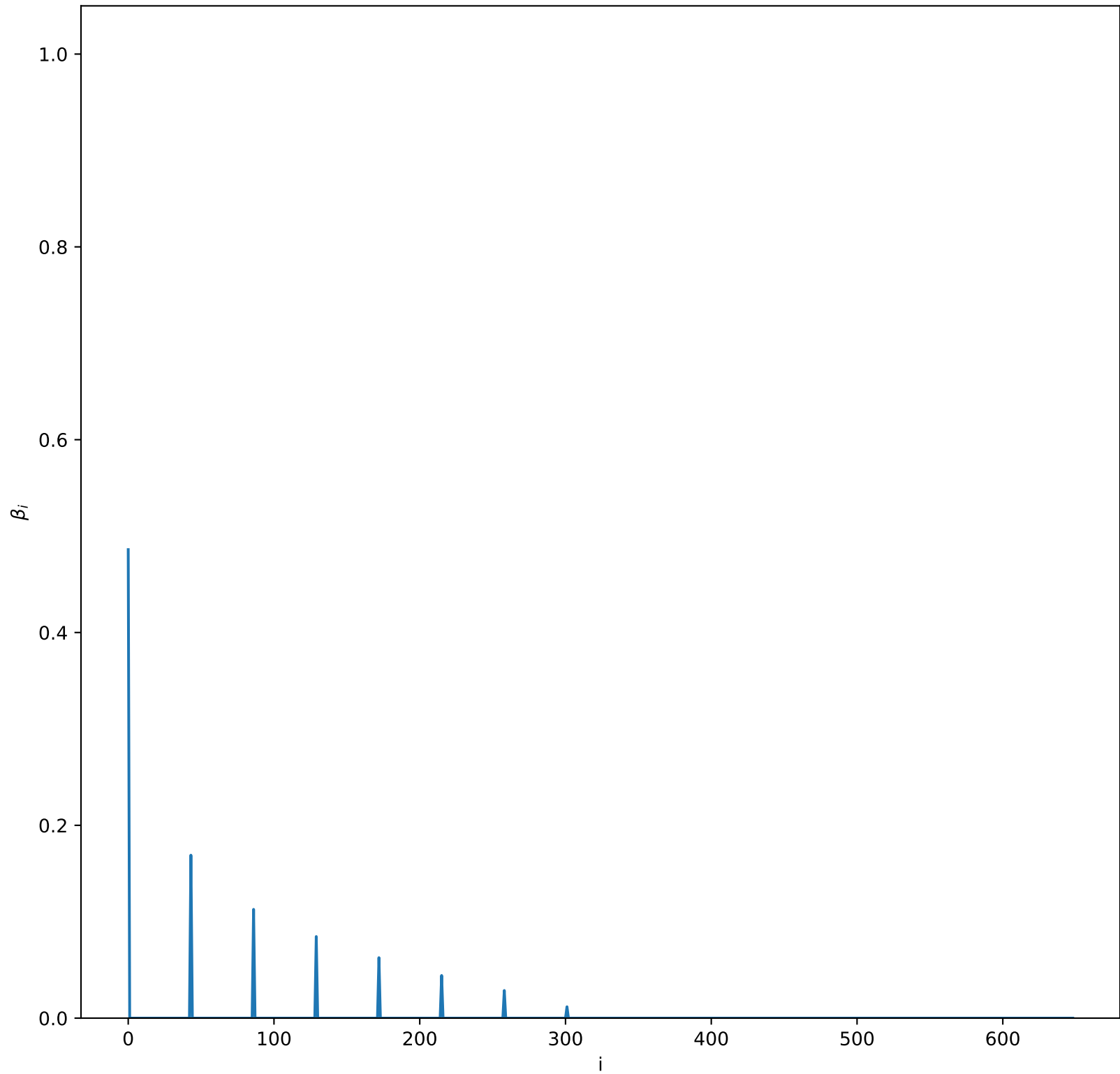
$\mu = 0.95$



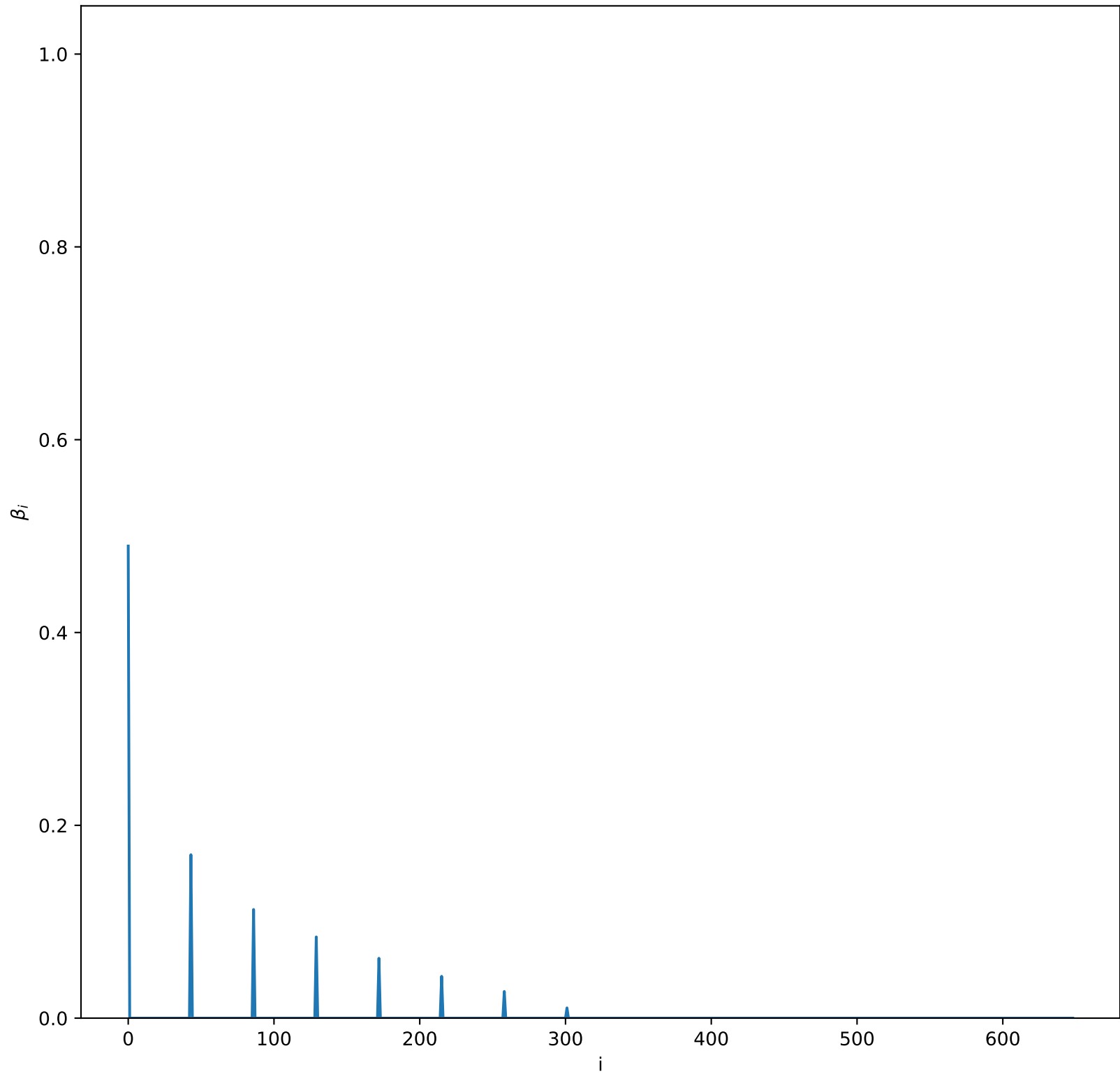
$\mu = 0.96$



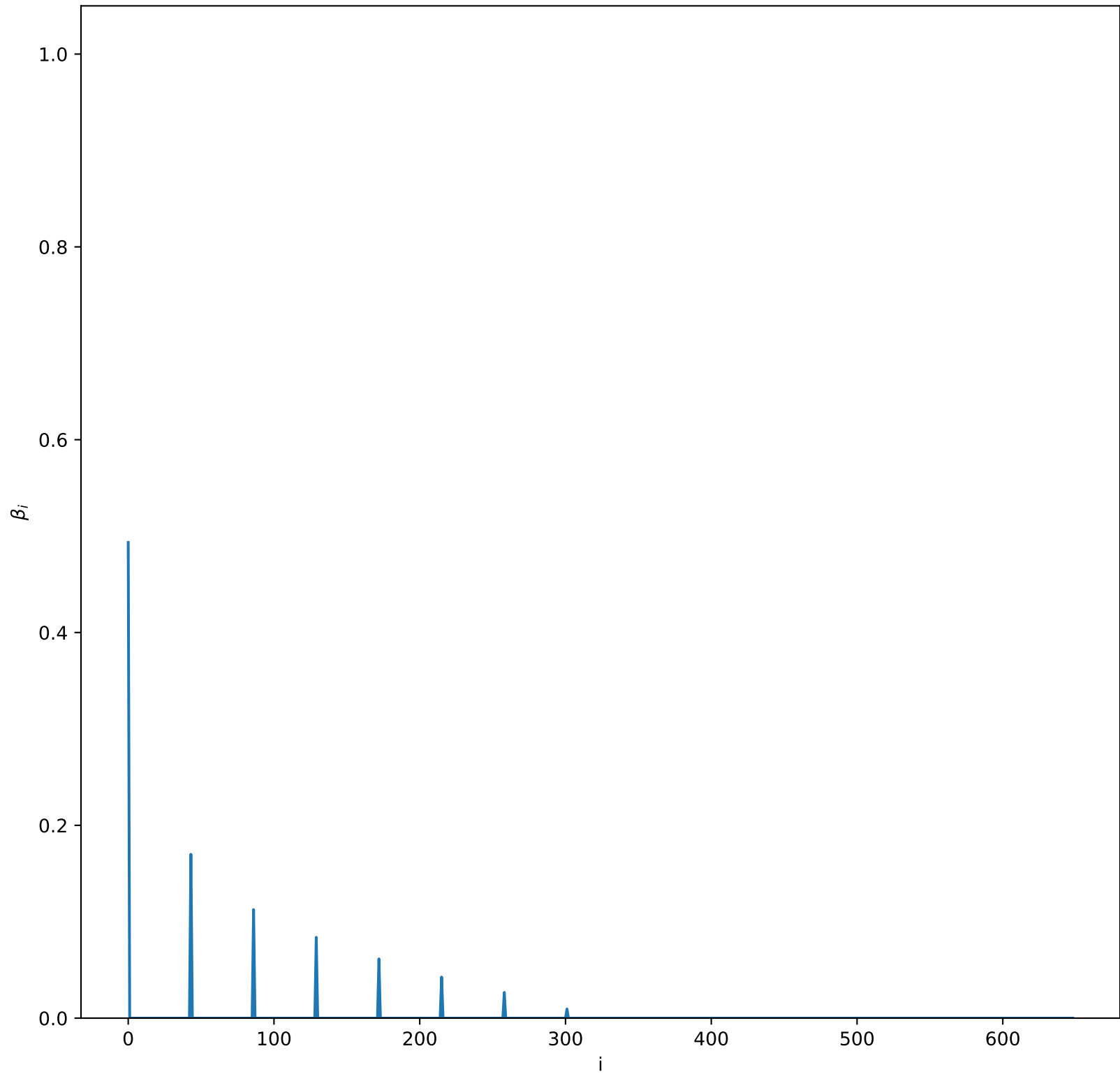
$\mu = 0.97$



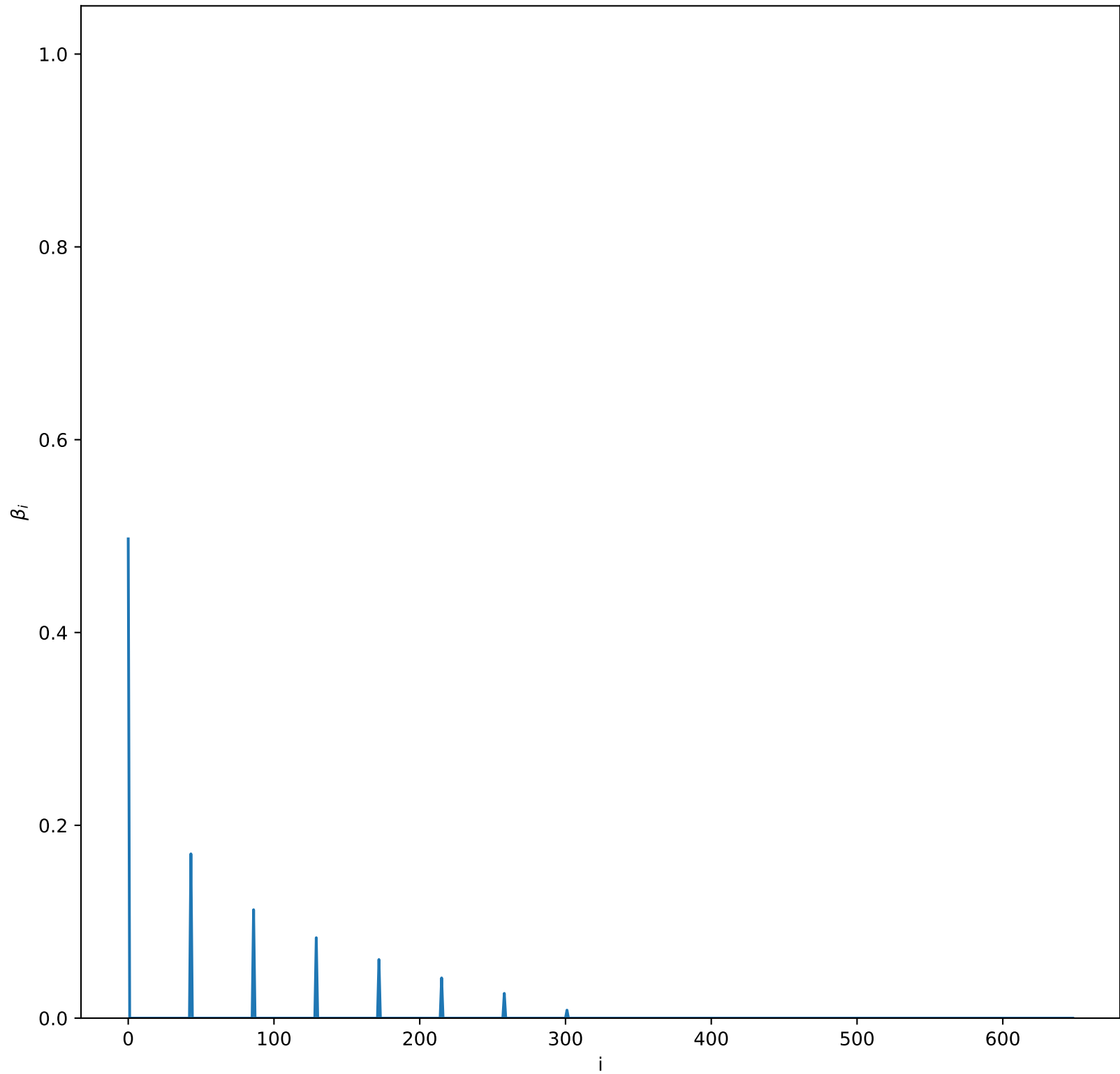
$\mu = 0.98$



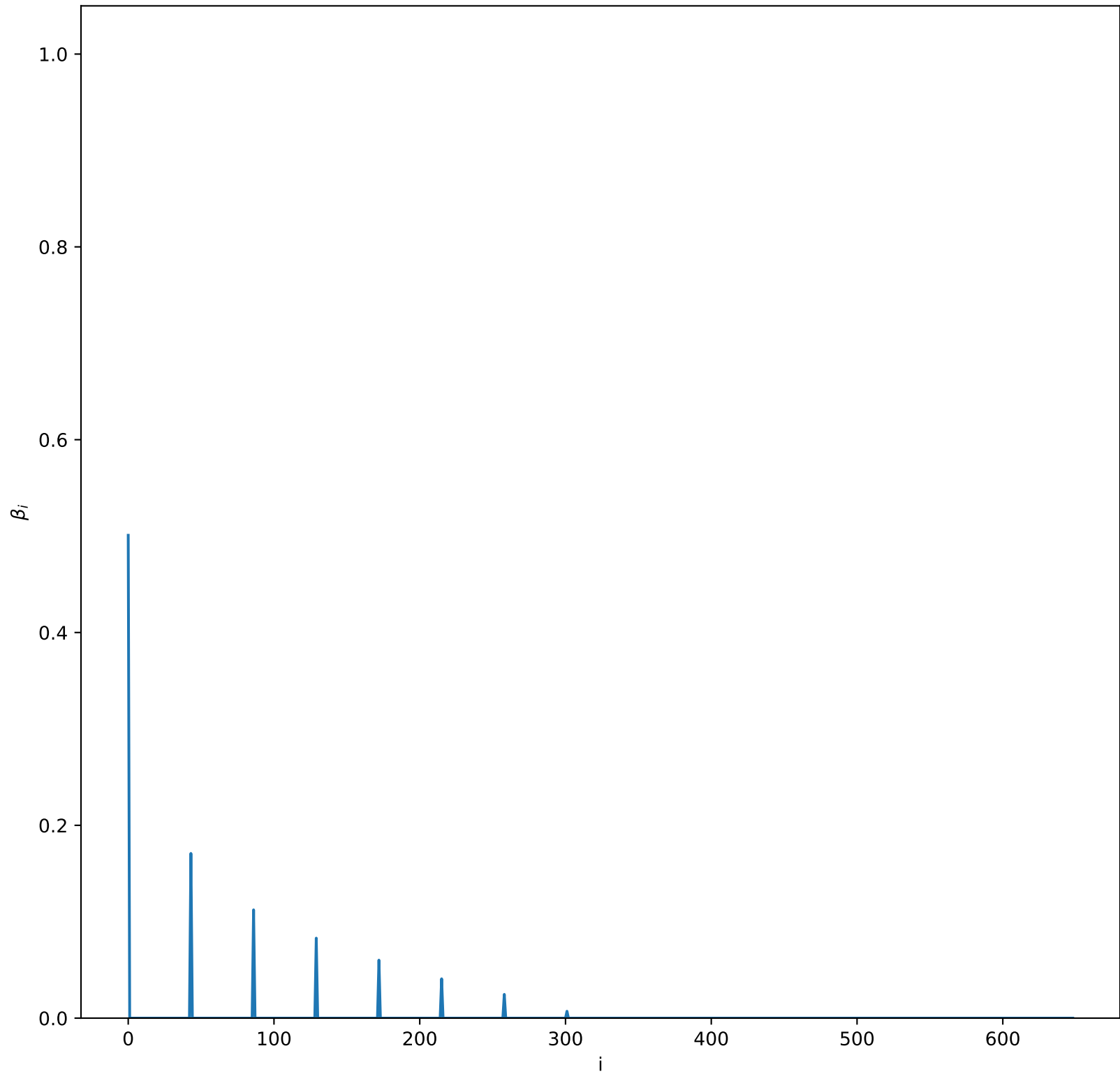
$\mu = 0.99$



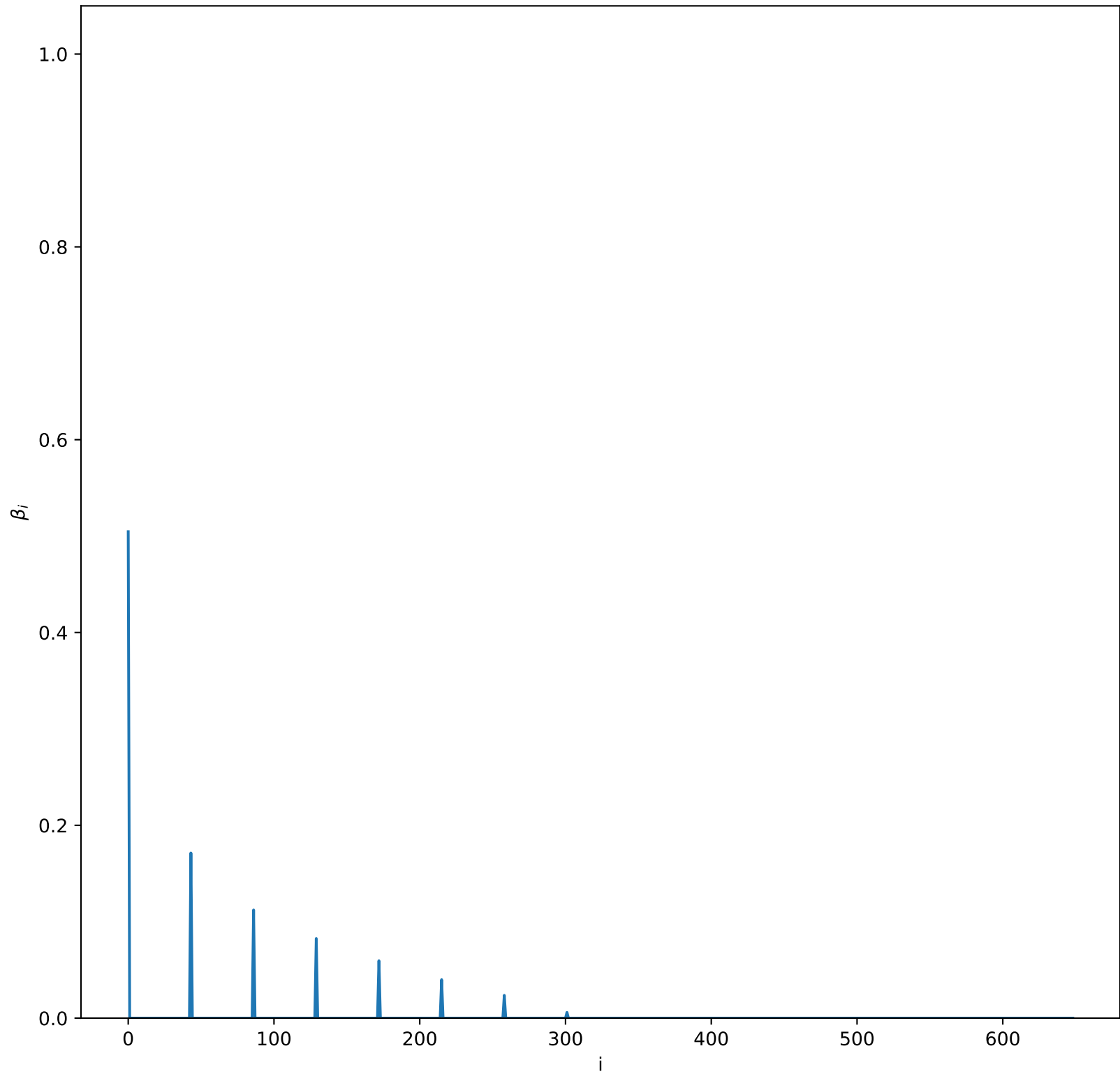
$\mu = 1.00$



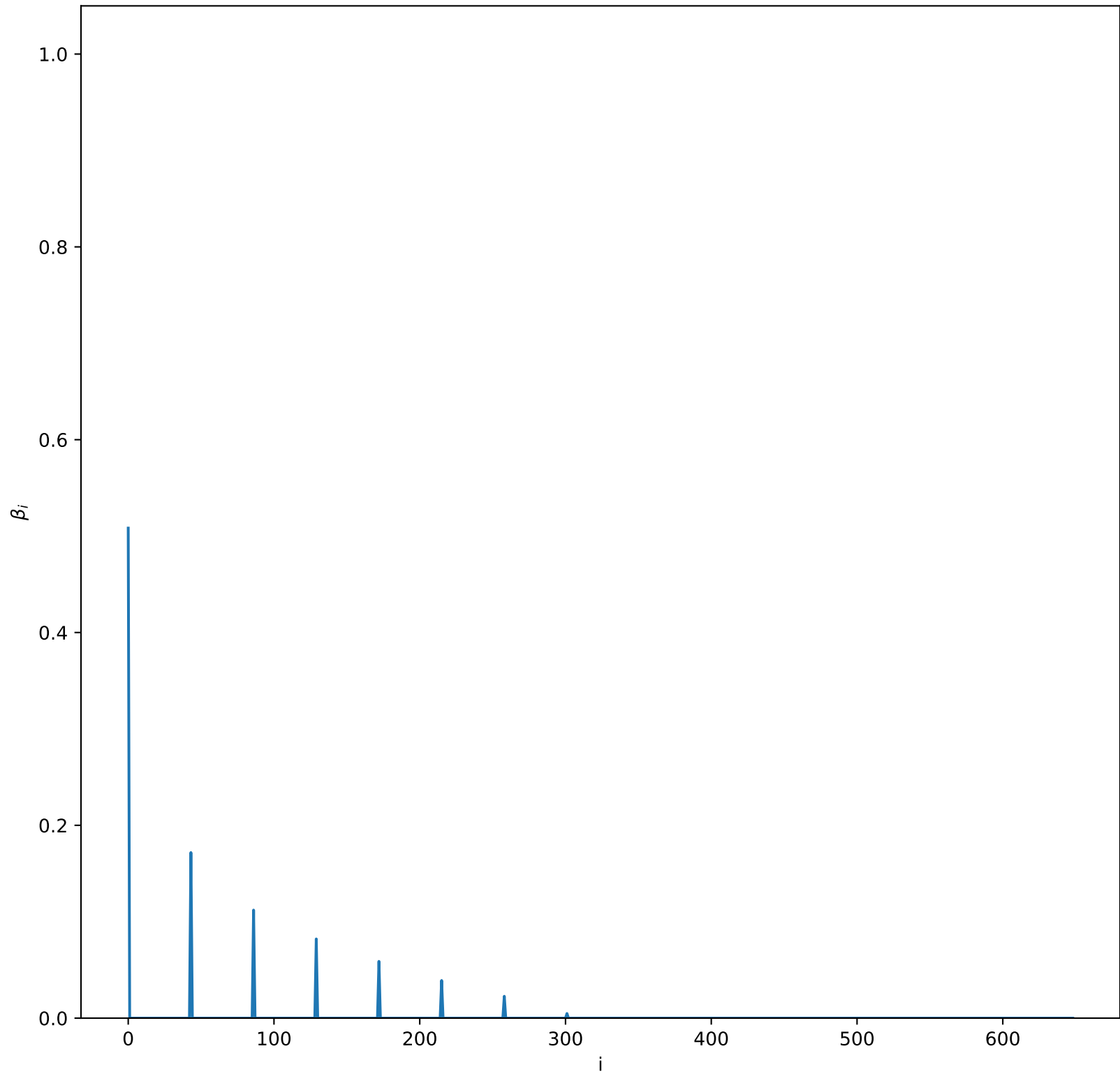
$\mu = 1.01$



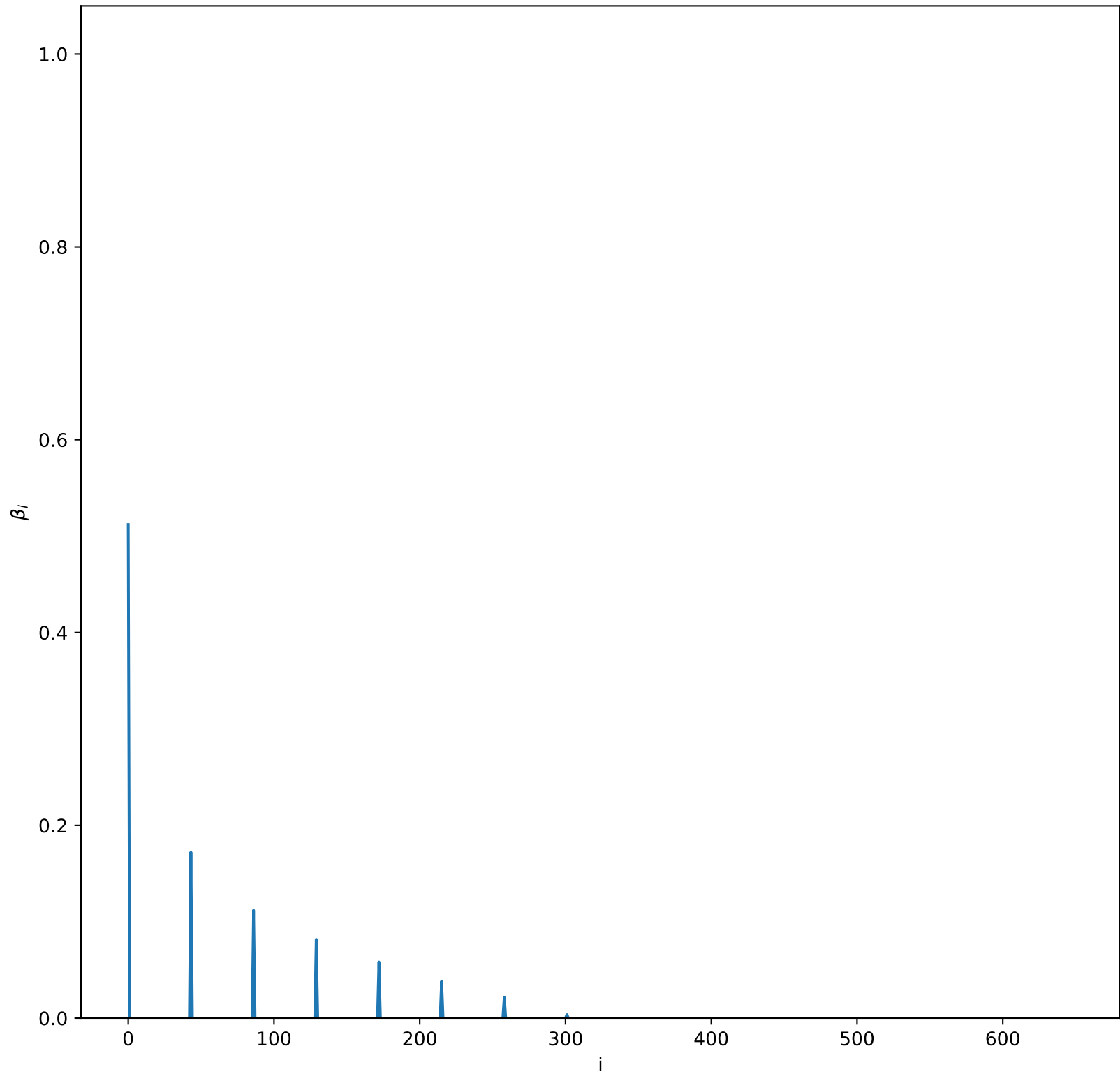
$\mu = 1.02$



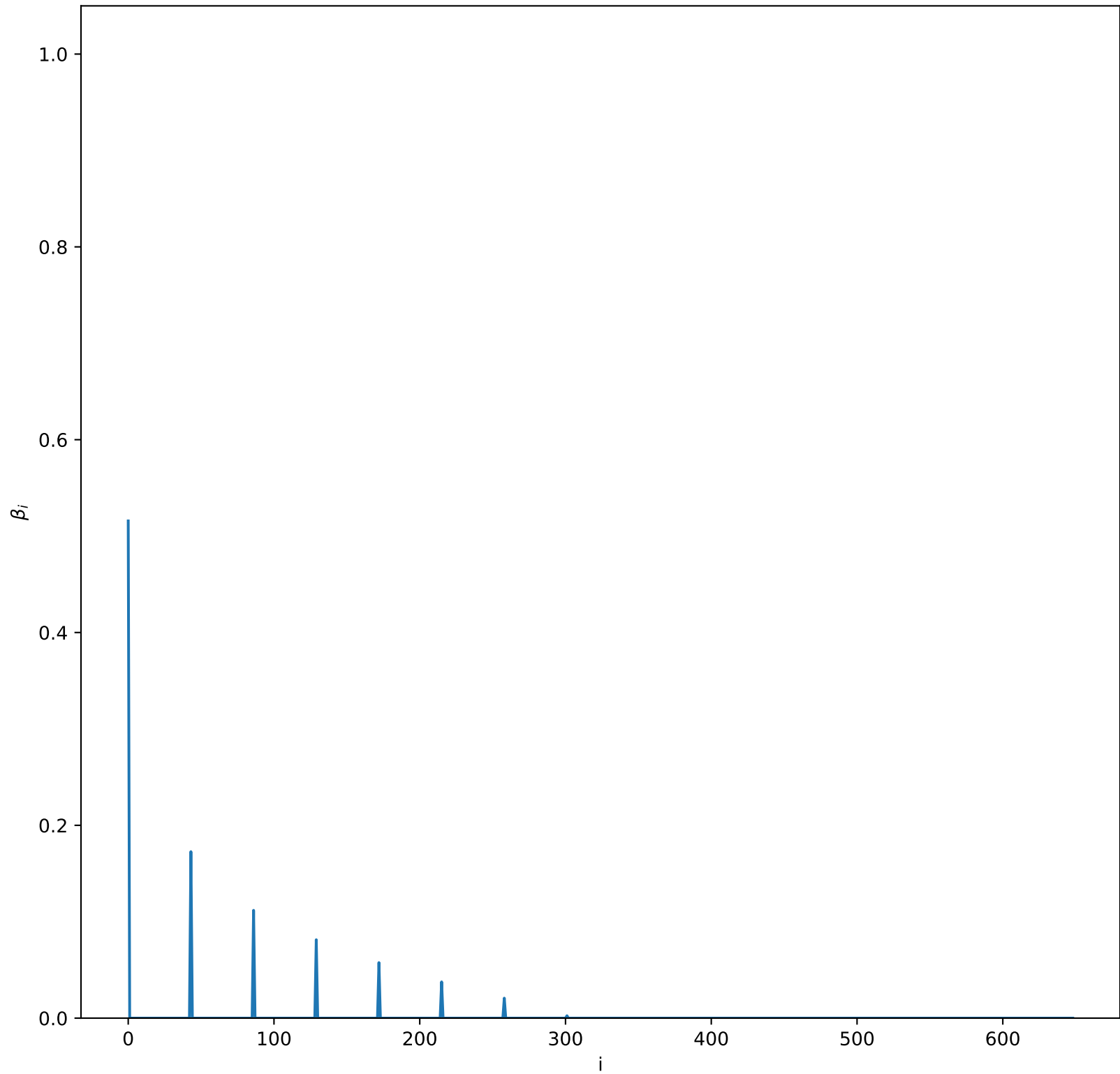
$\mu = 1.03$



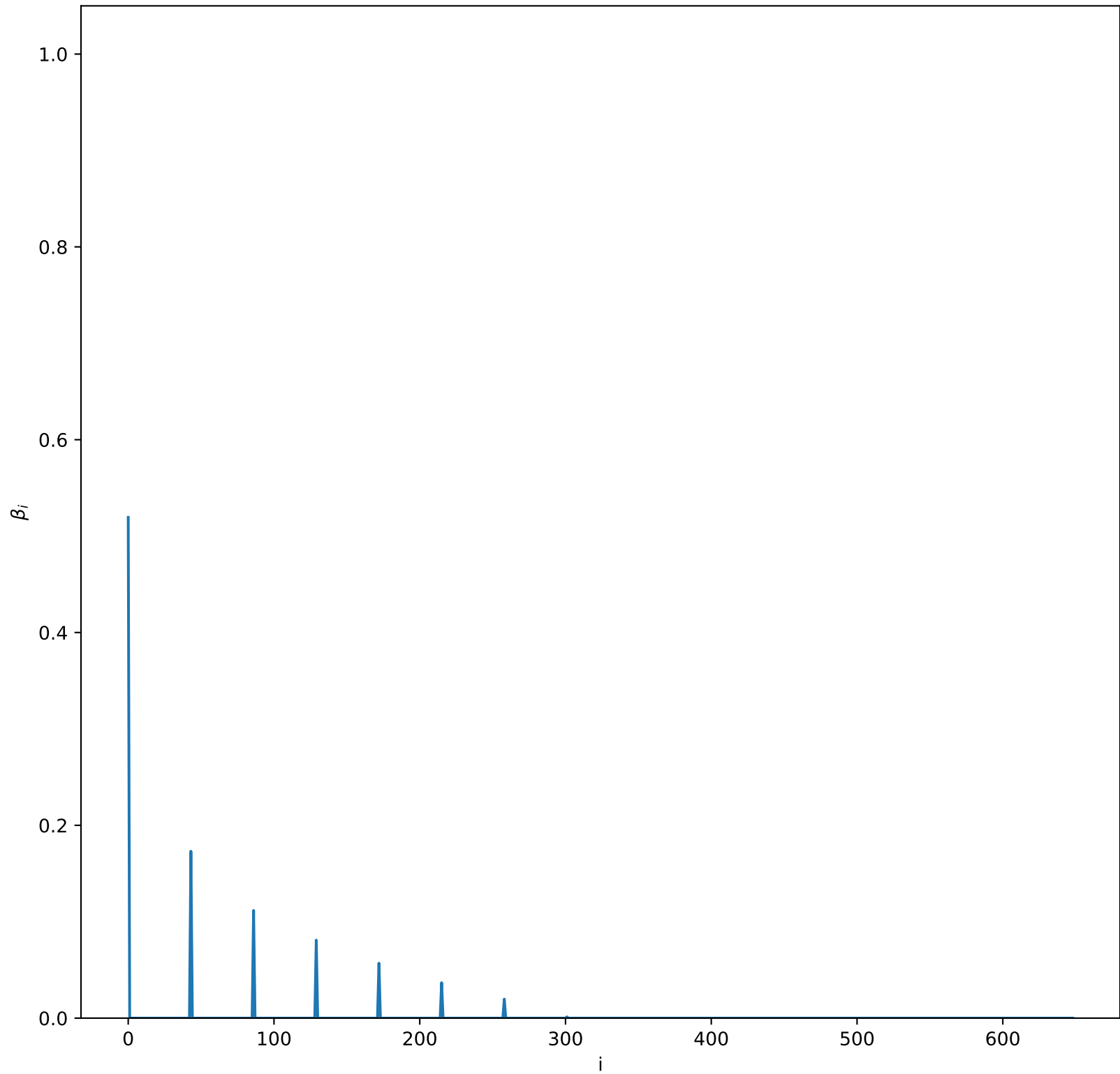
$\mu = 1.04$



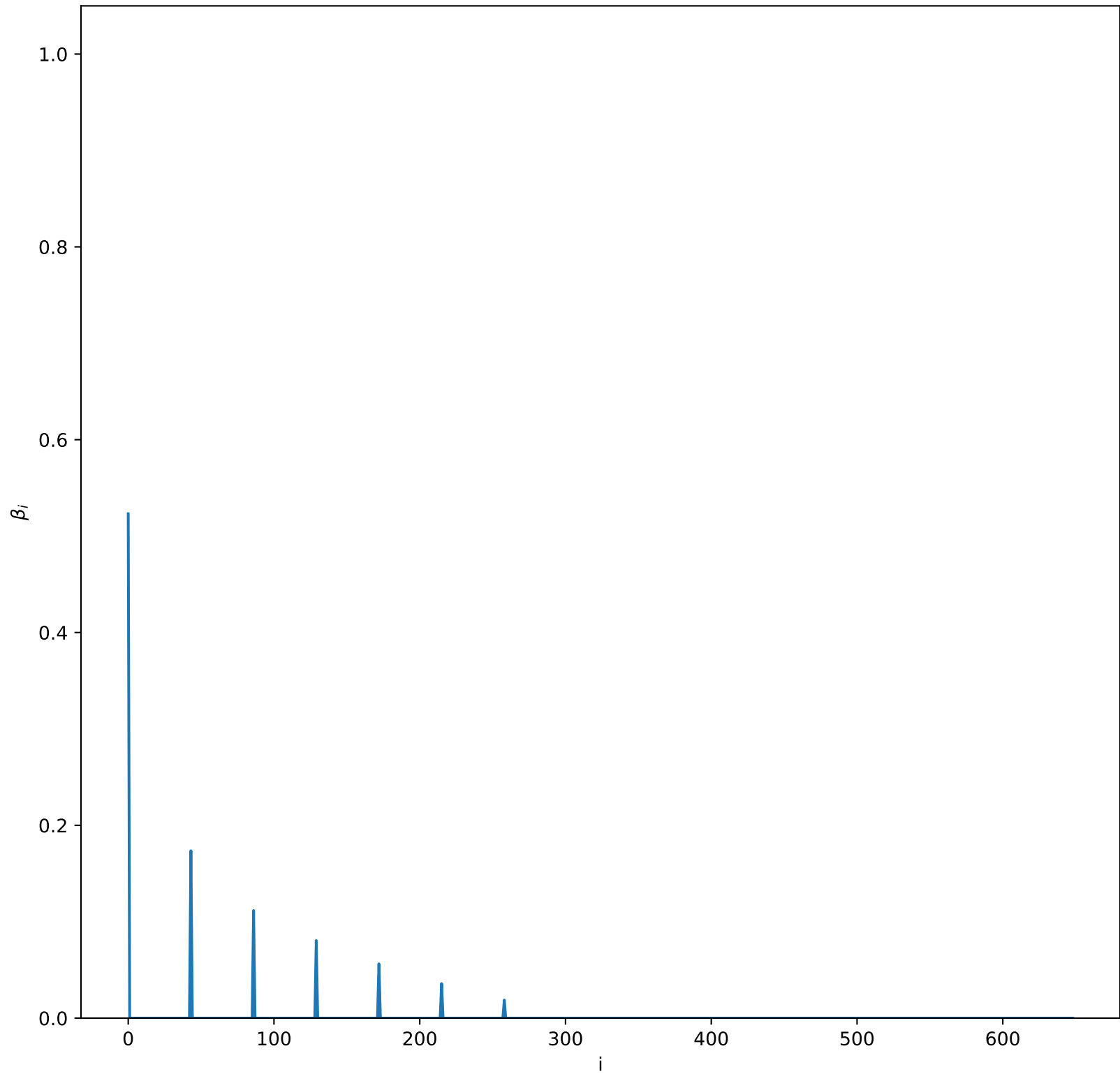
$\mu = 1.05$



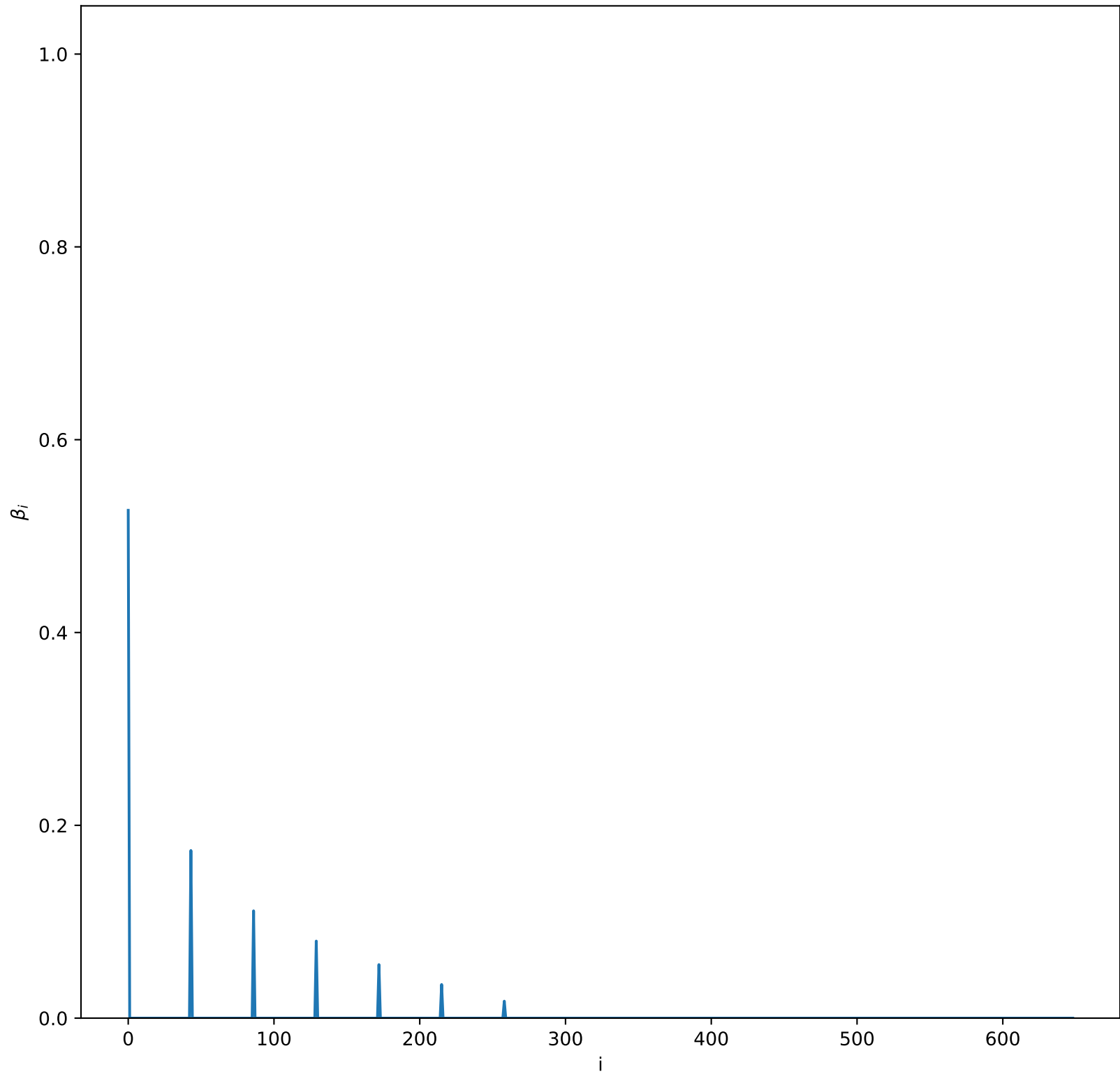
$\mu = 1.06$



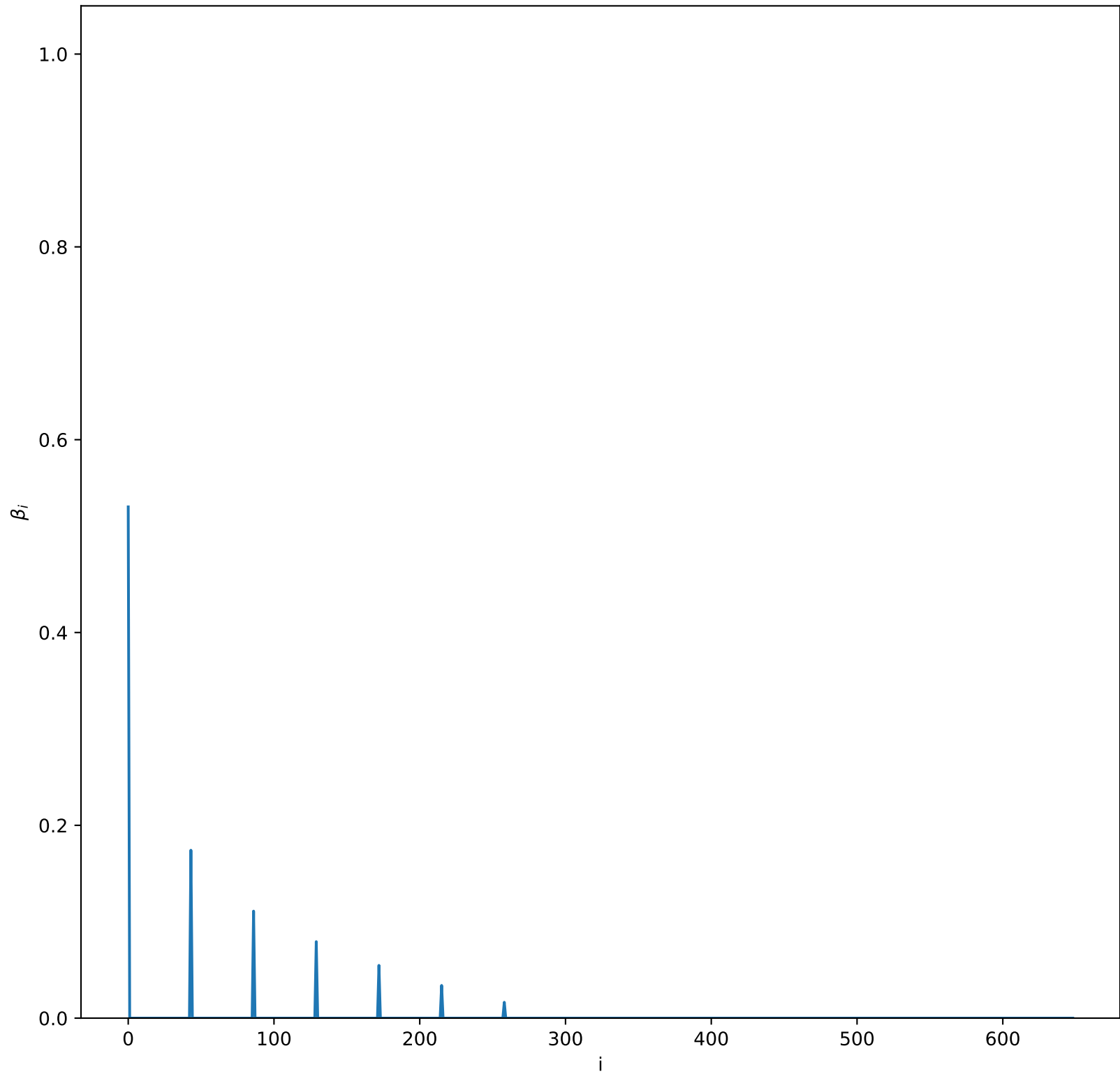
$\mu = 1.07$



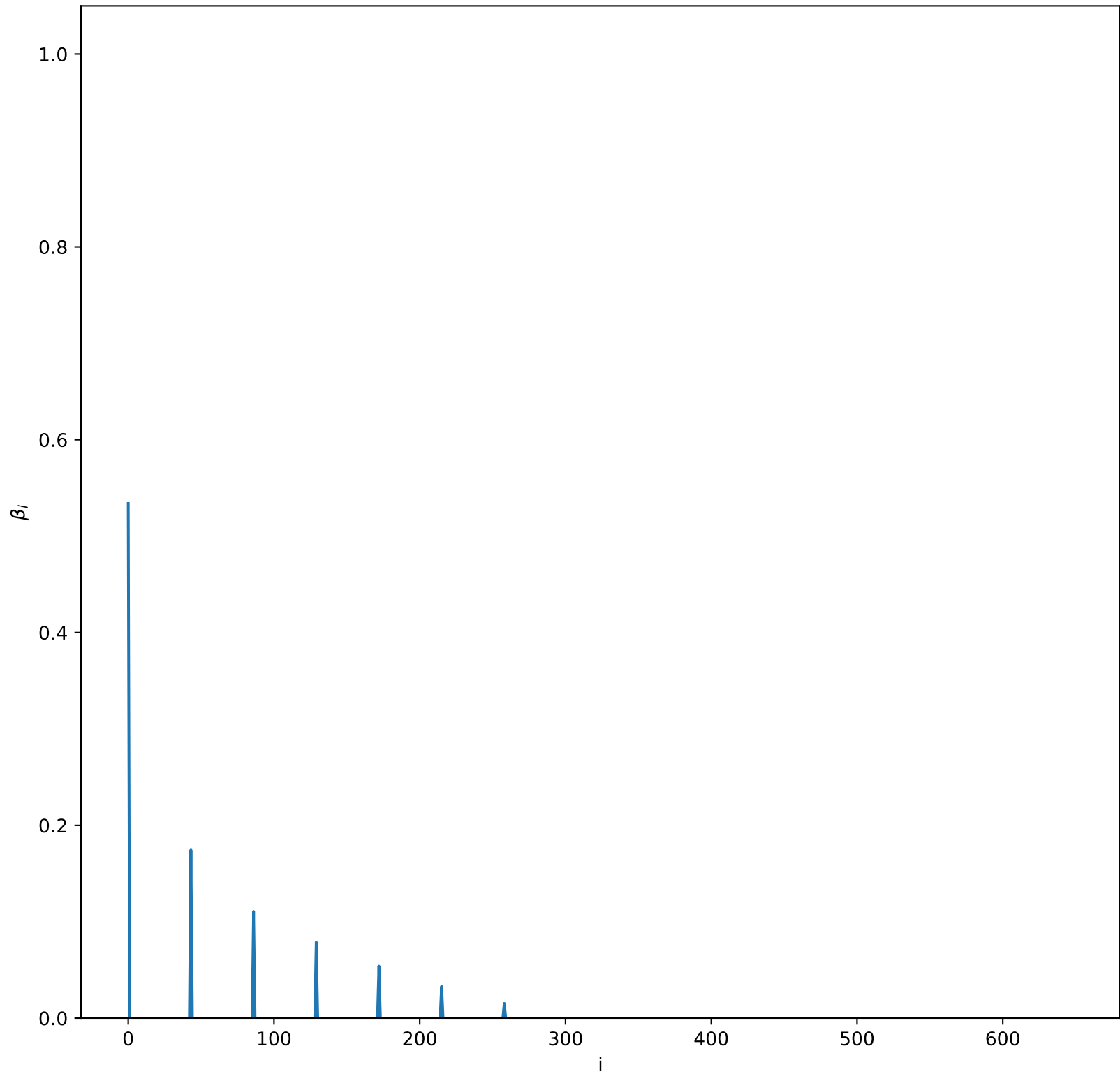
$\mu = 1.08$



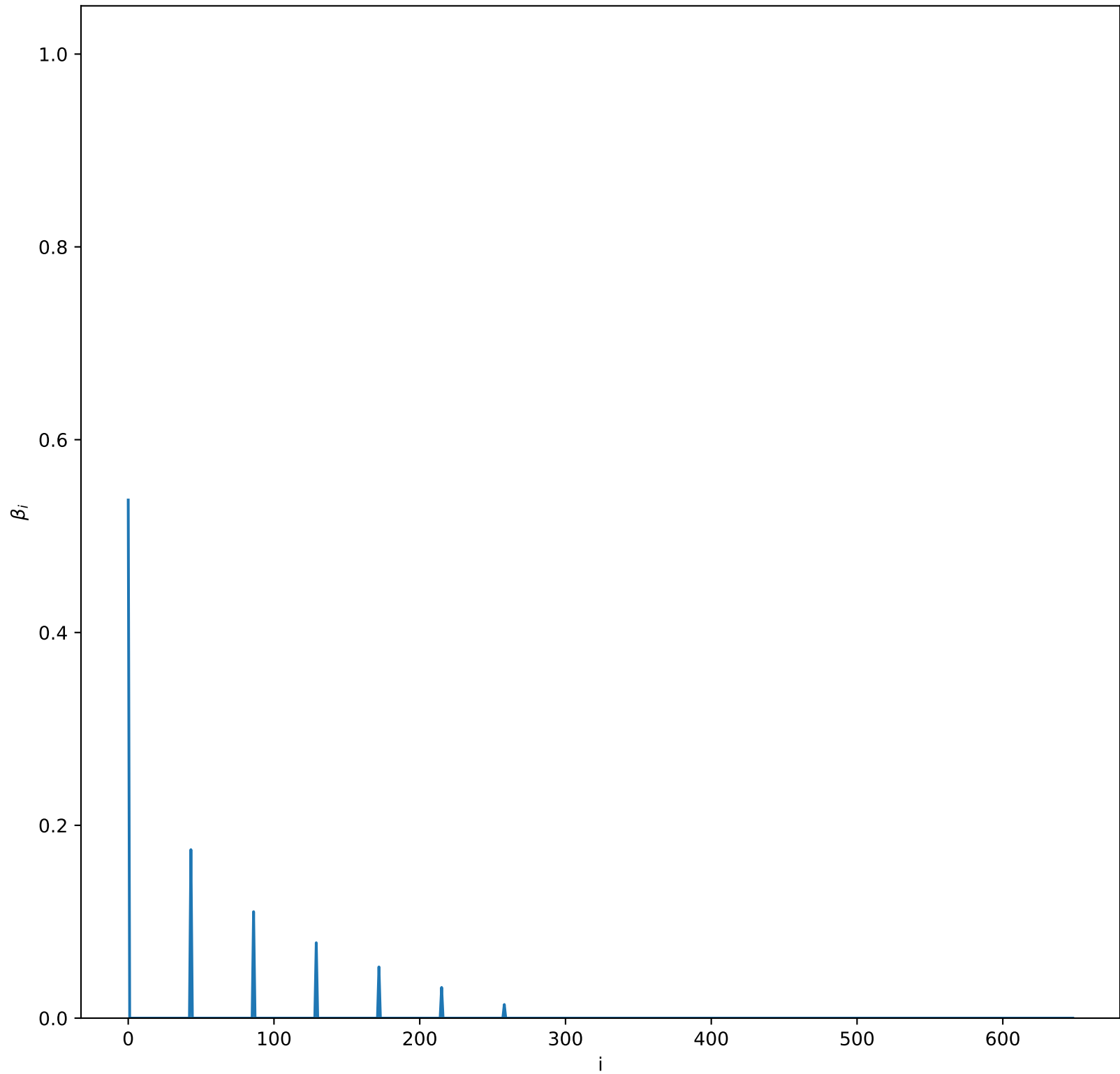
$\mu = 1.09$



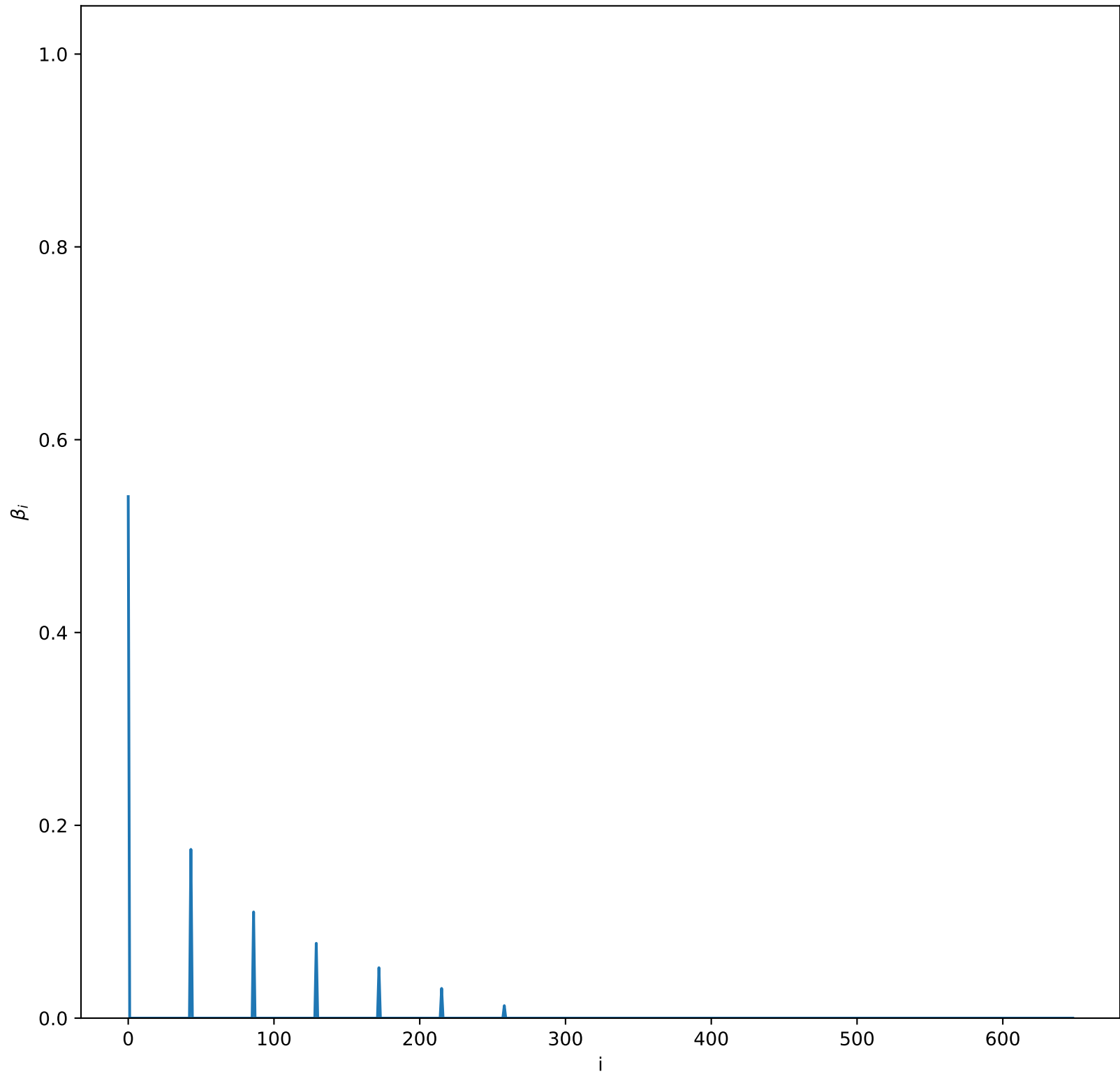
$\mu = 1.10$



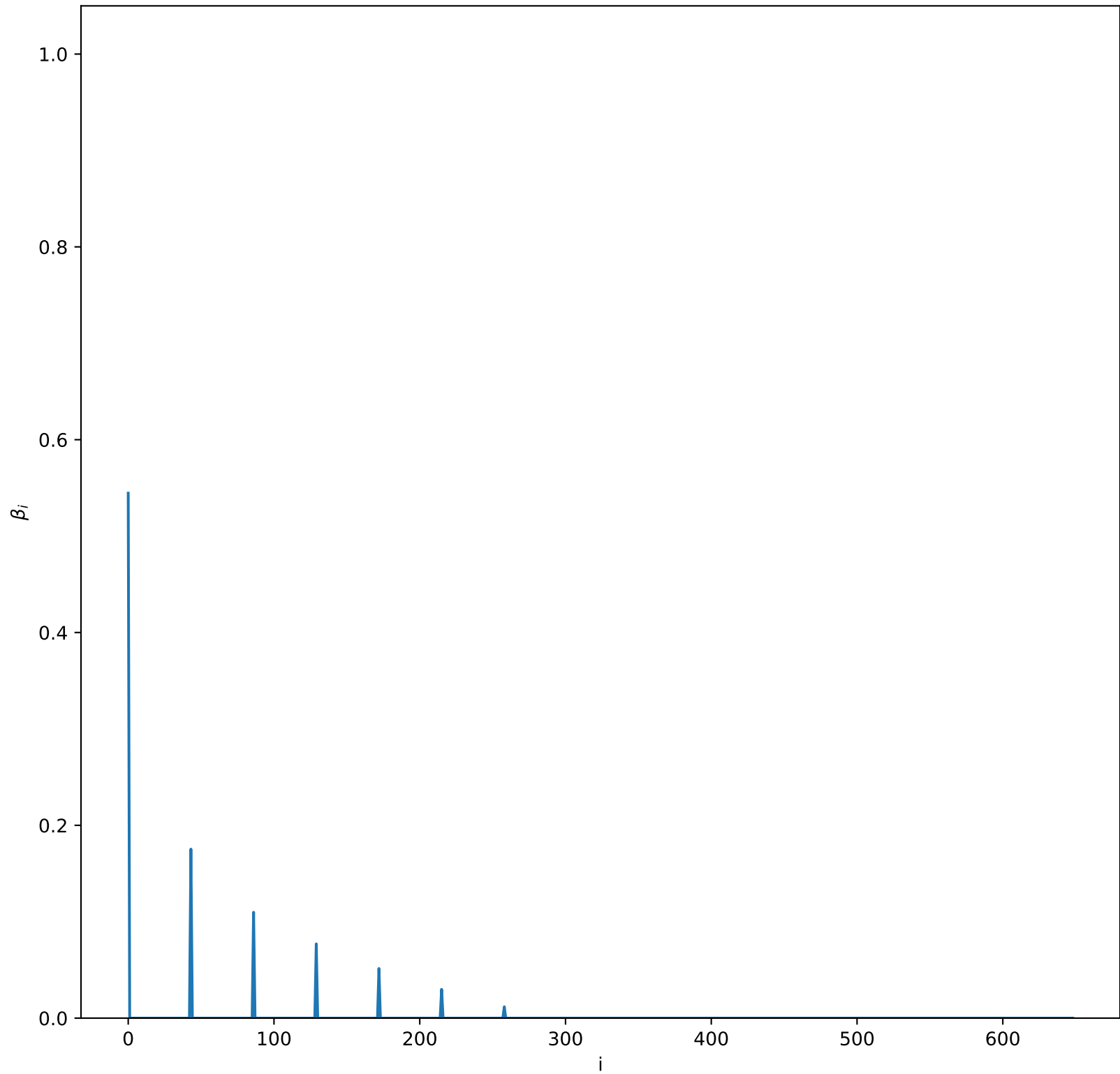
$\mu = 1.11$



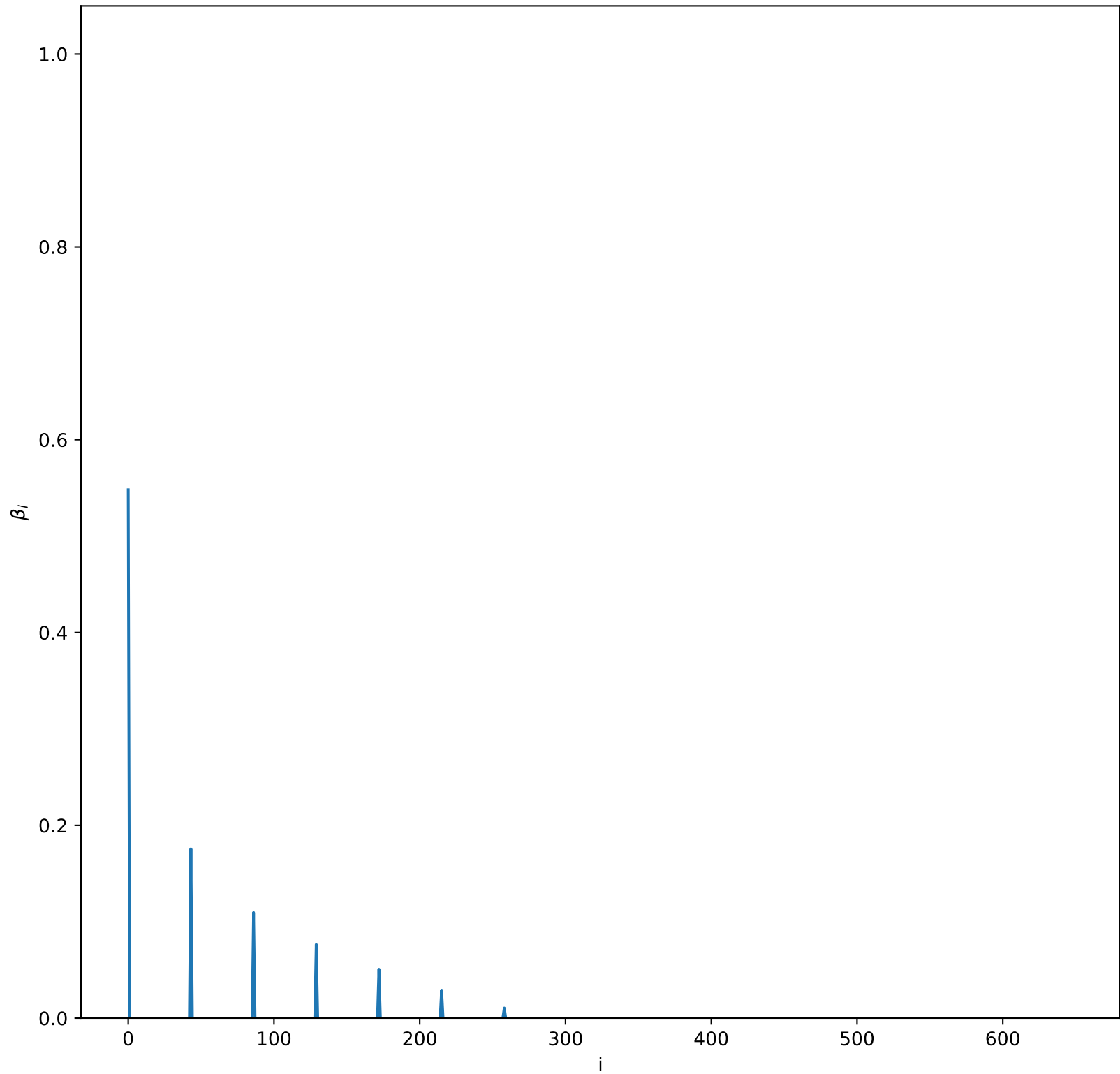
$\mu = 1.12$



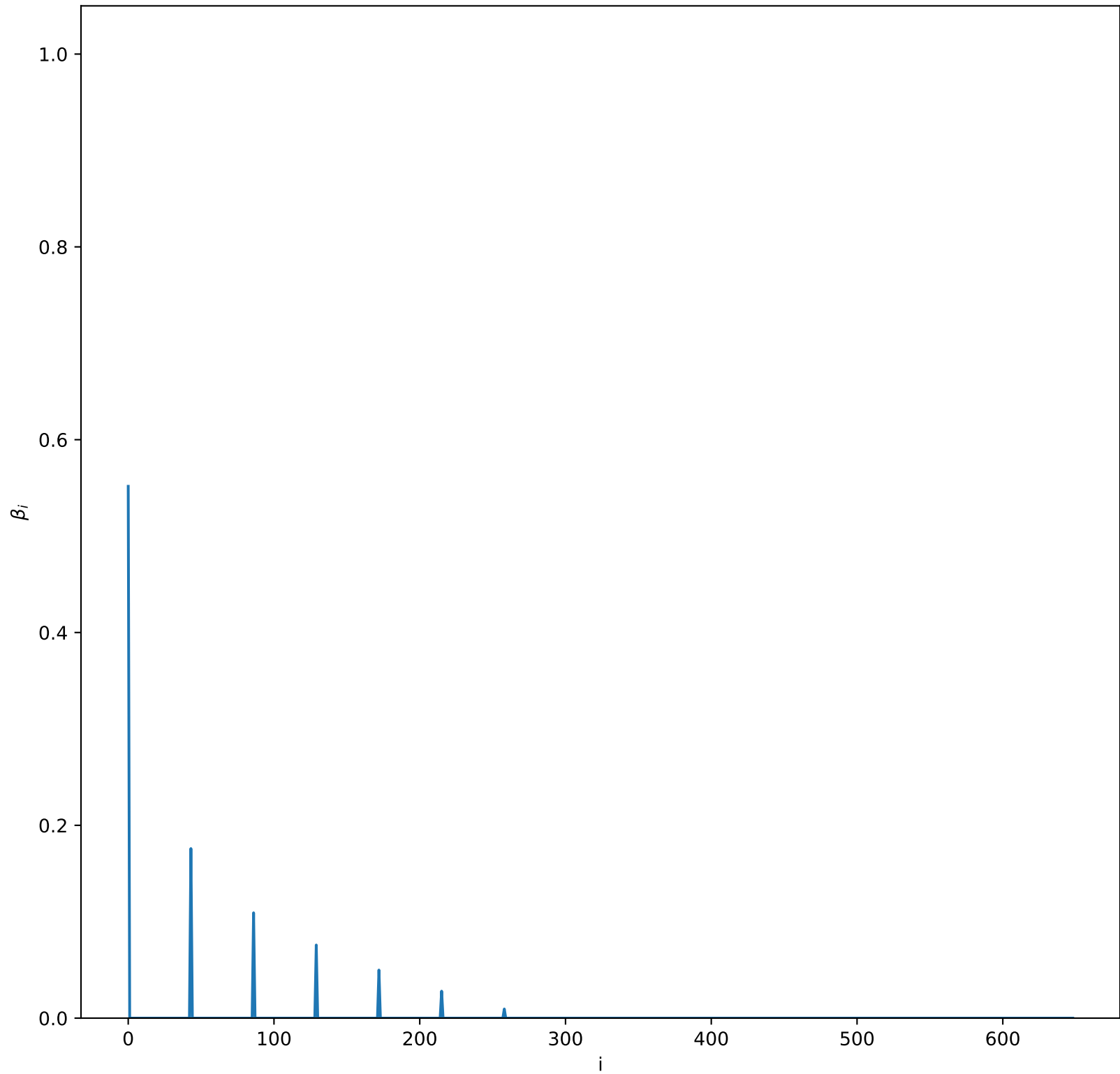
$\mu = 1.13$



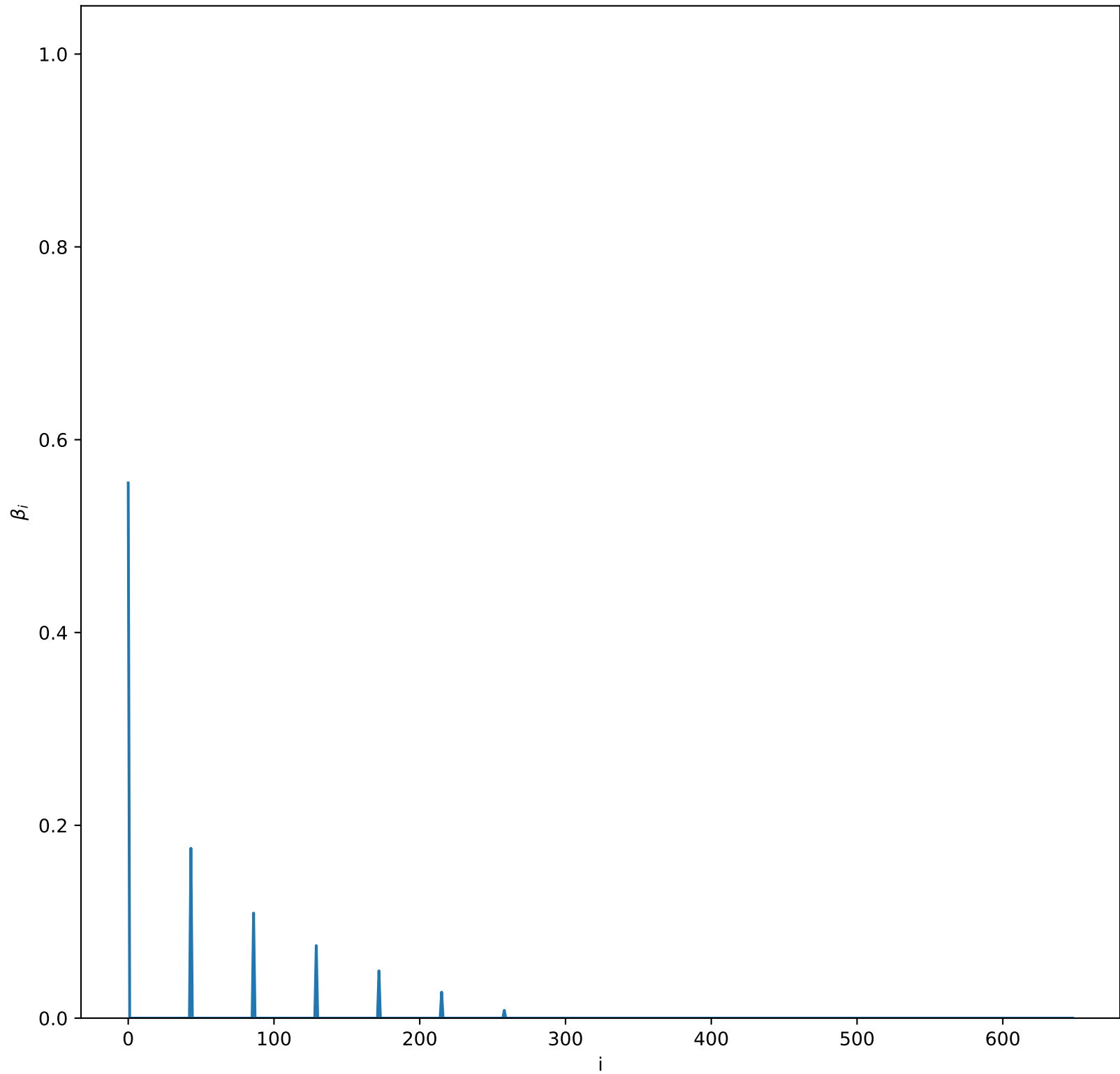
$\mu = 1.14$



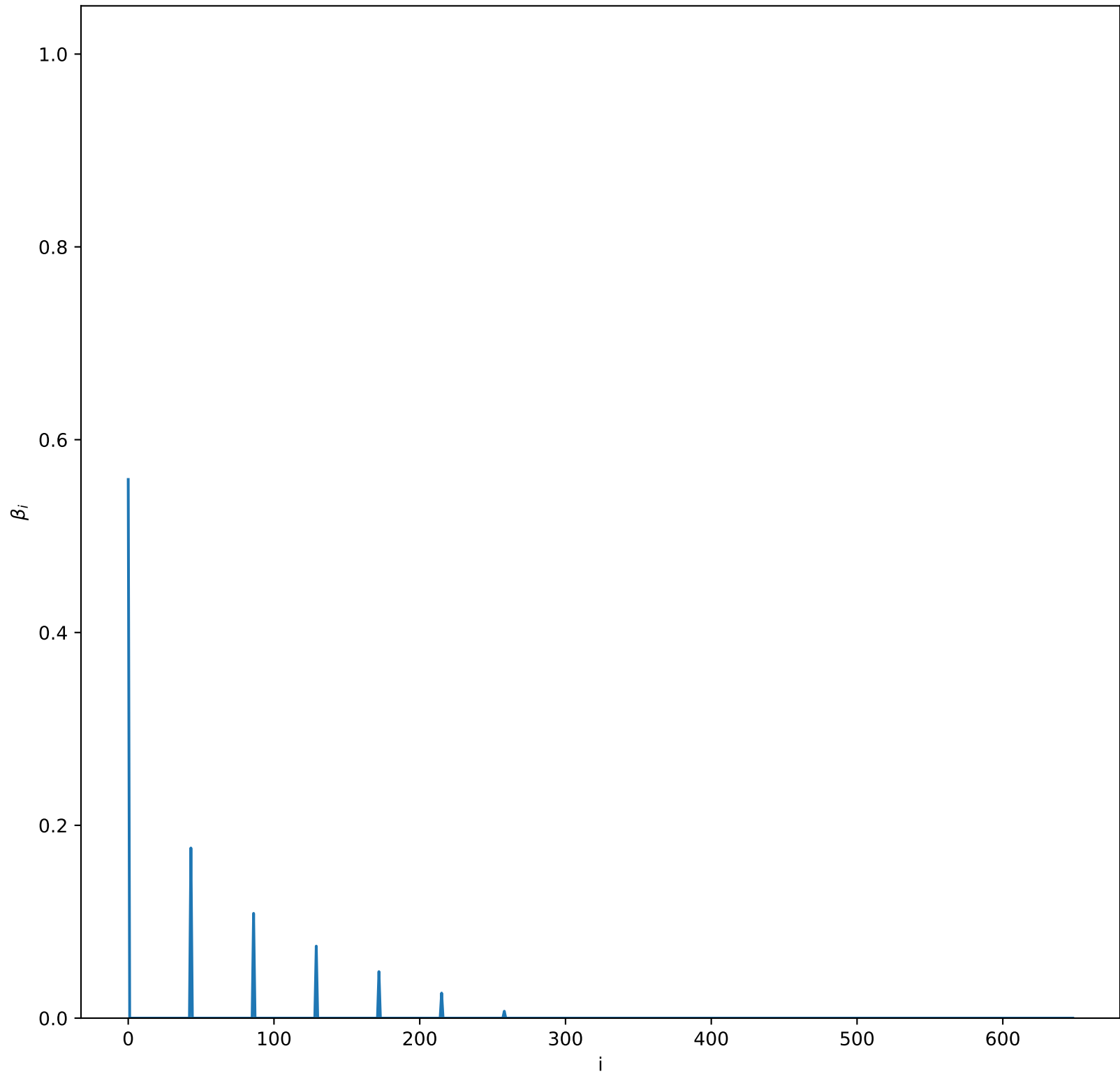
$\mu = 1.15$



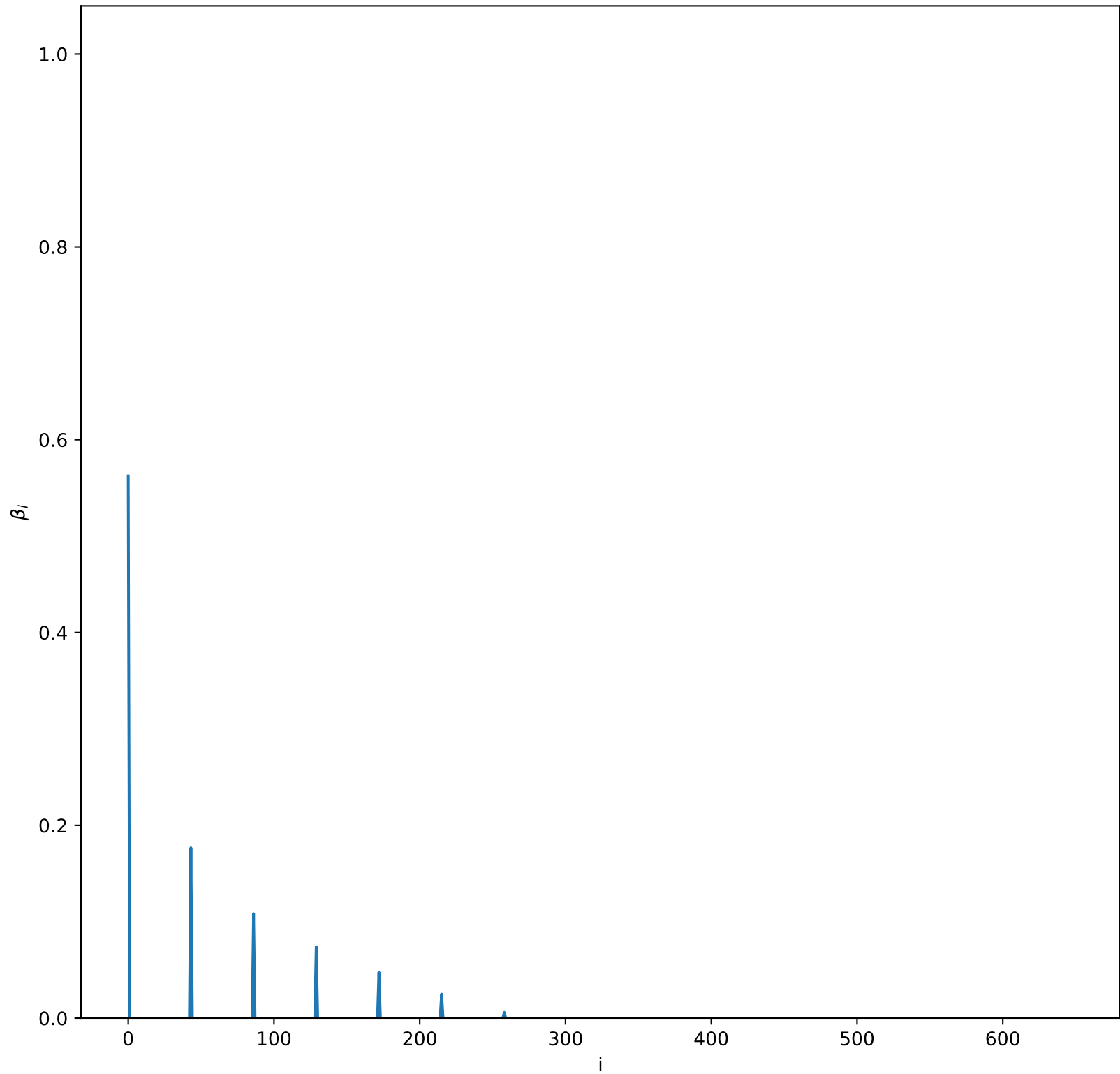
$\mu = 1.16$



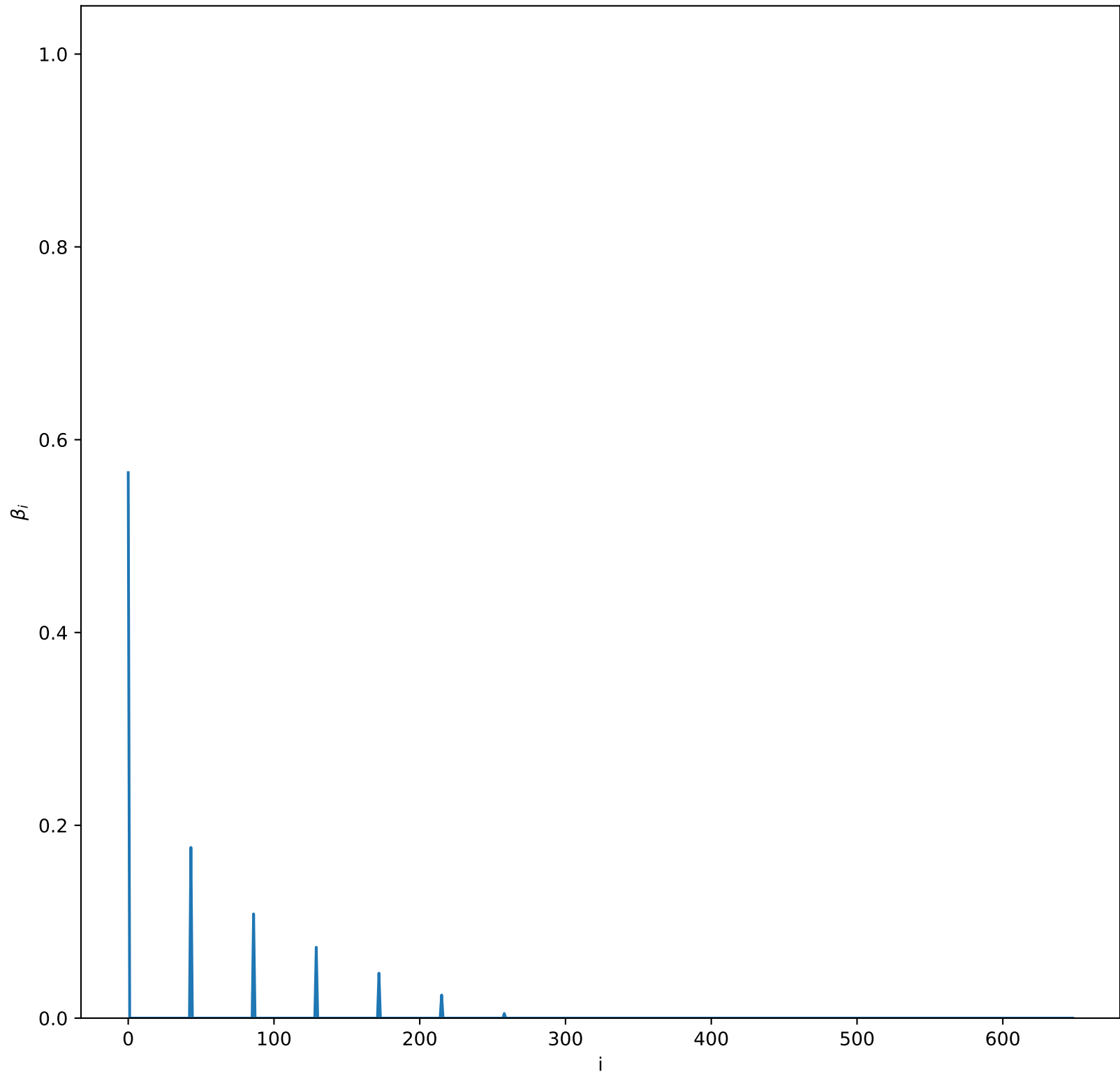
$\mu = 1.17$



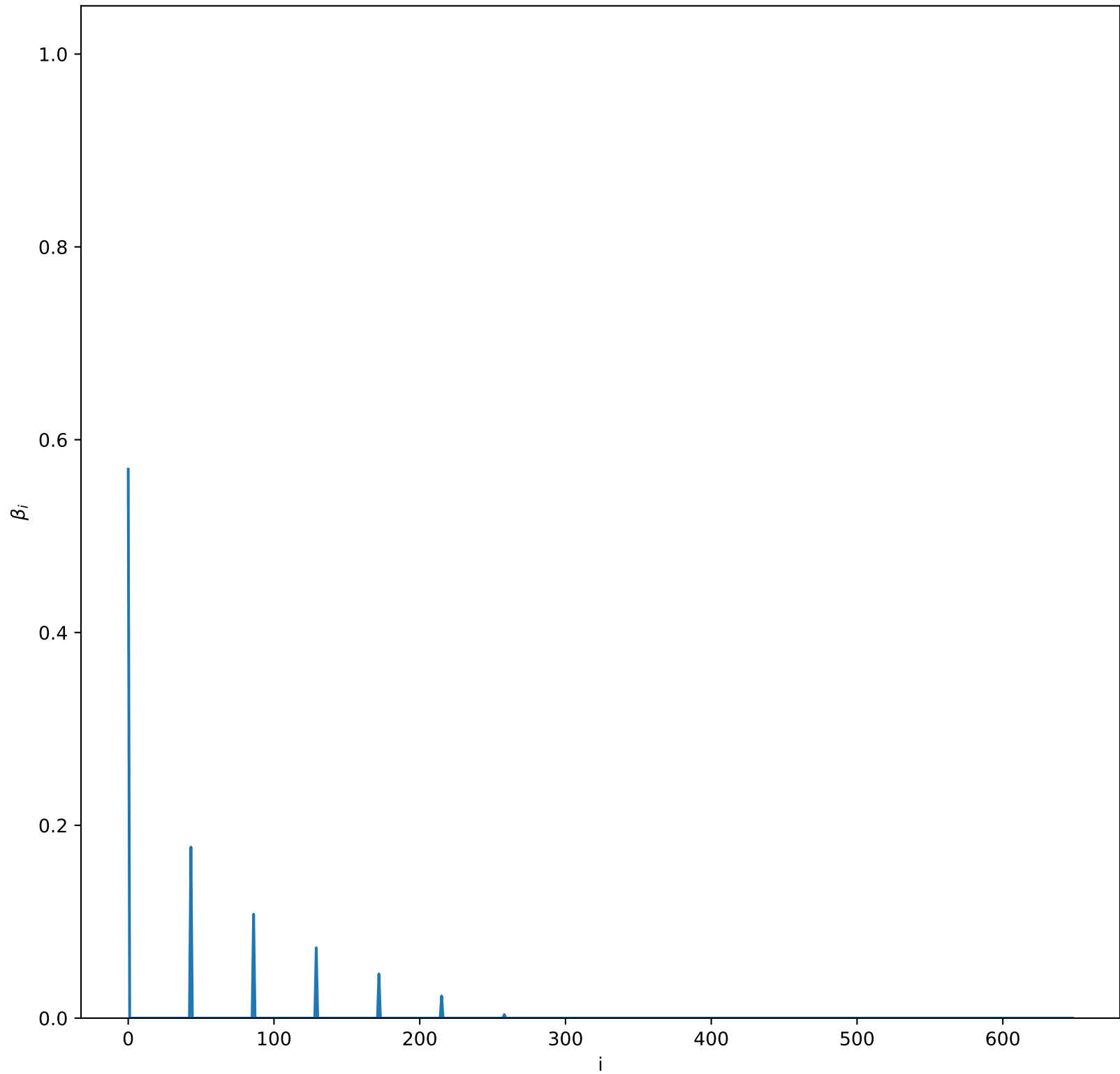
$\mu = 1.18$



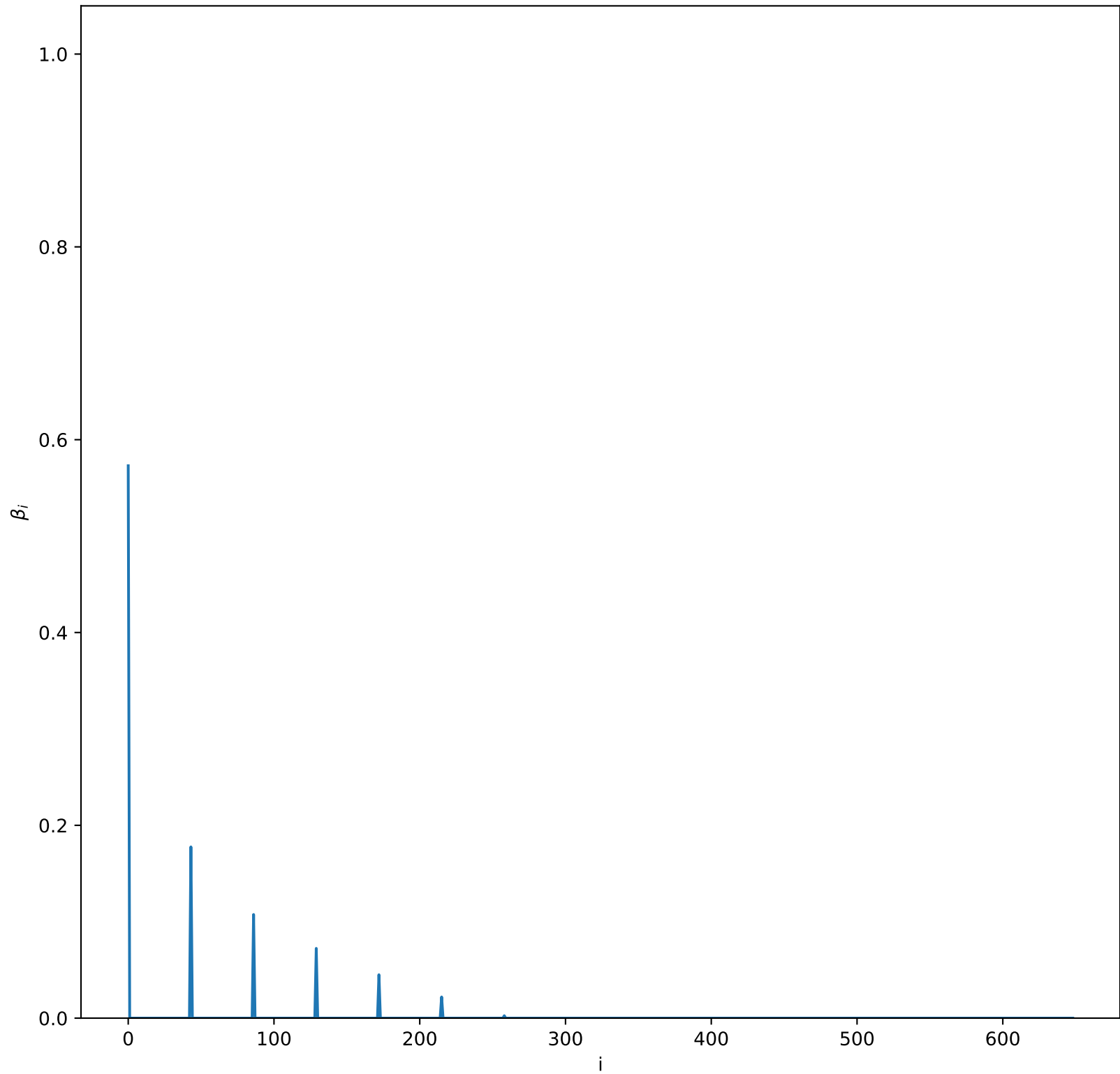
$\mu = 1.19$



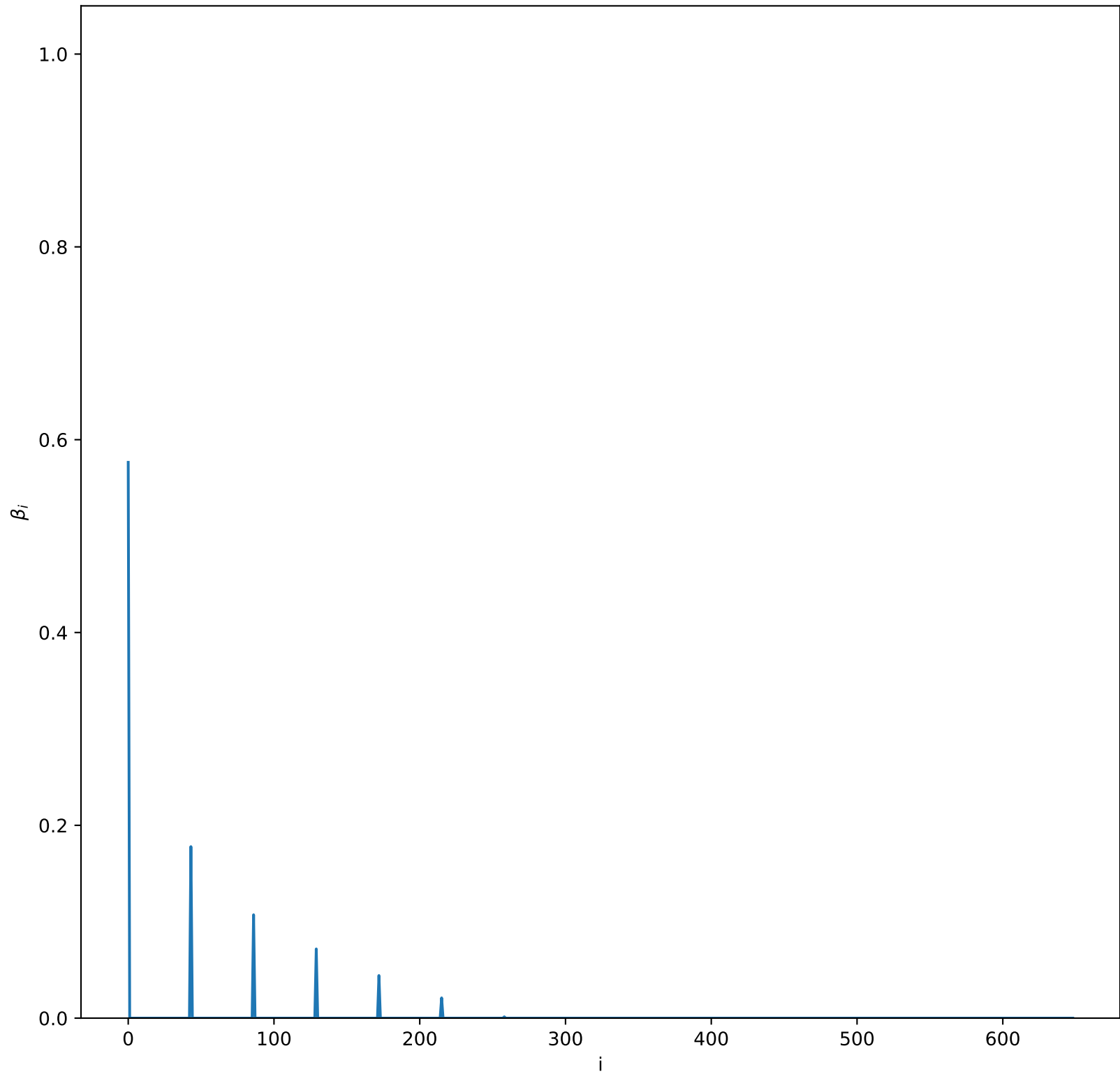
$\mu = 1.20$



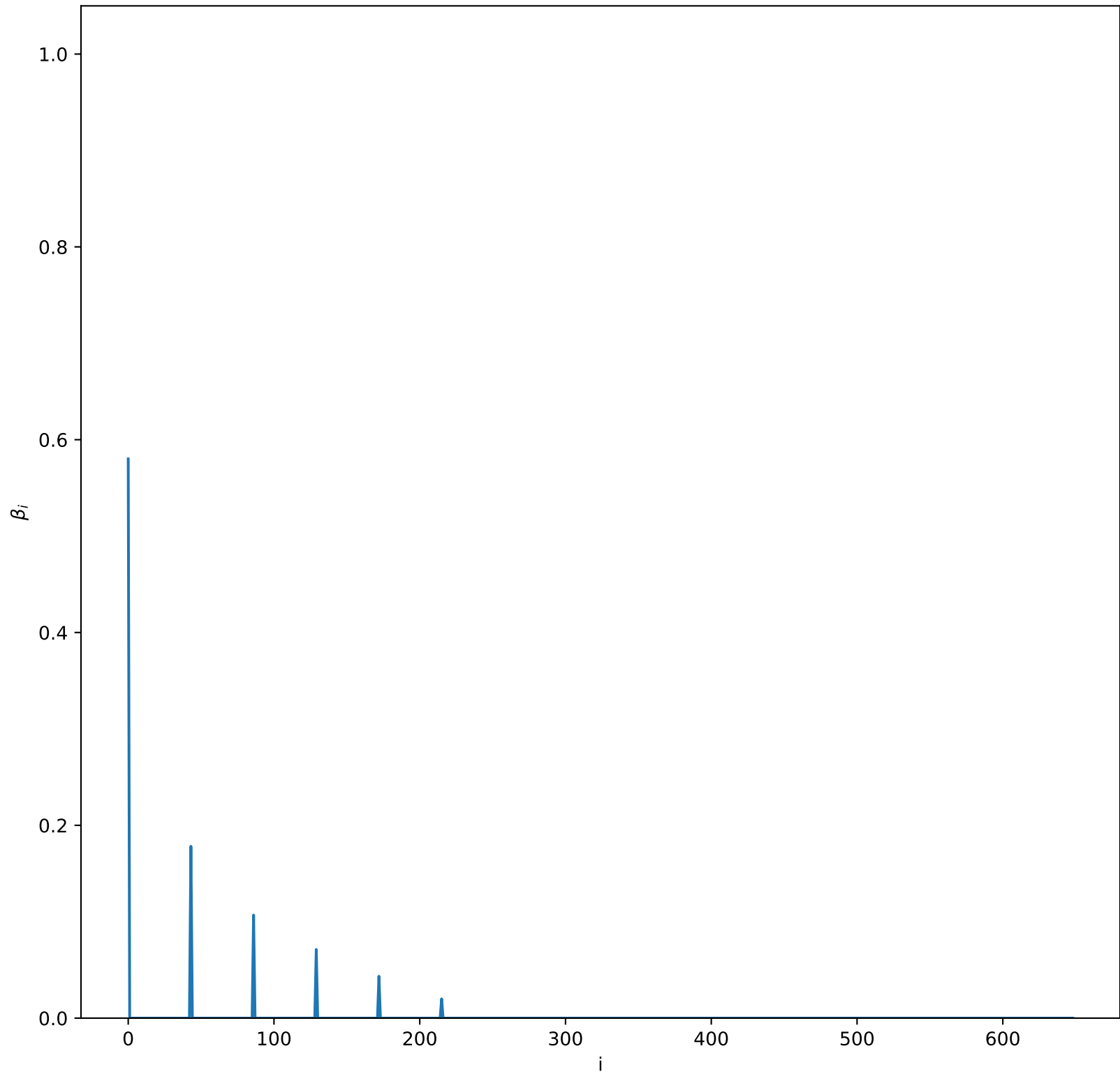
$\mu = 1.21$



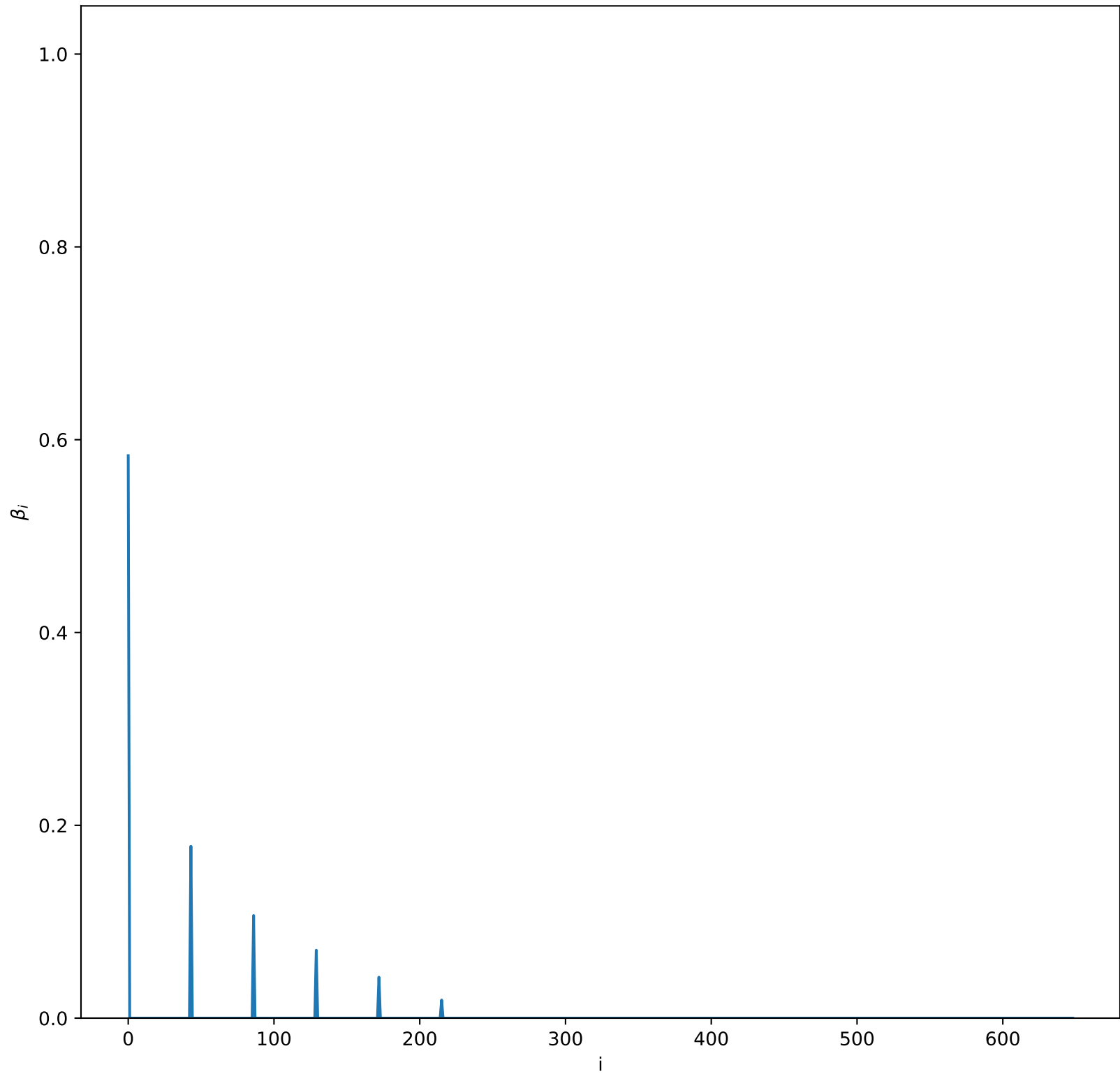
$\mu = 1.22$



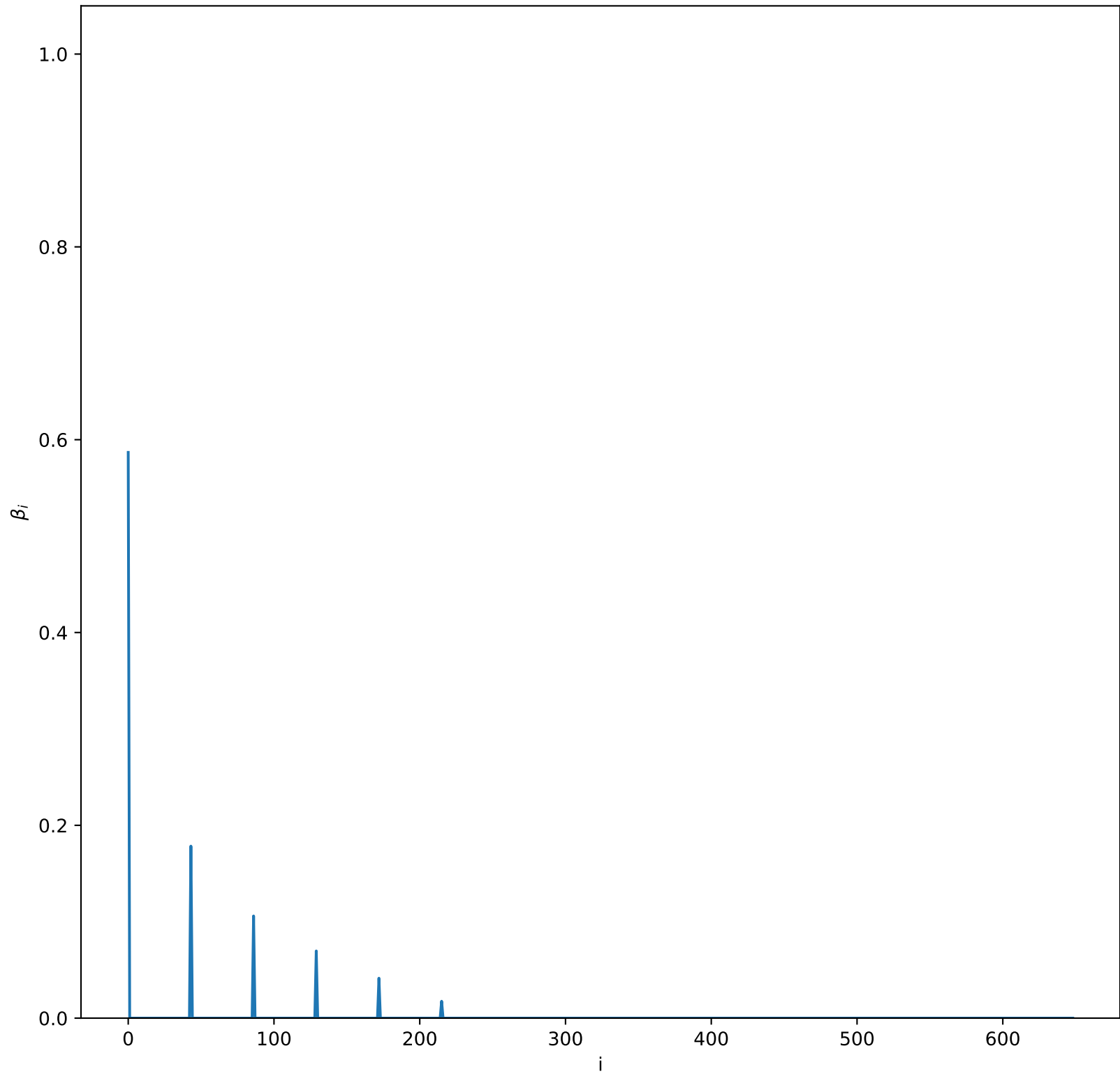
$\mu = 1.23$



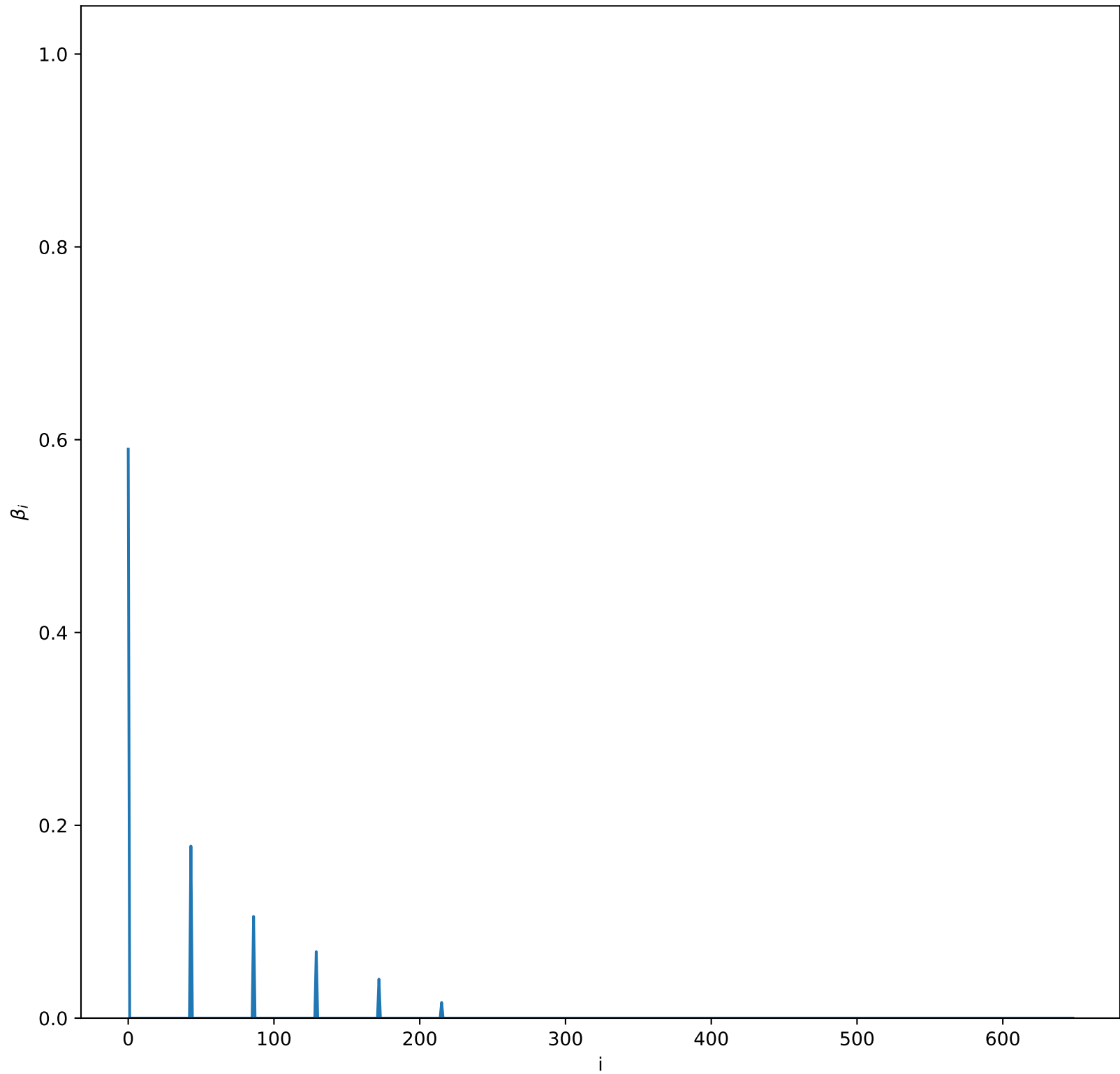
$\mu = 1.24$



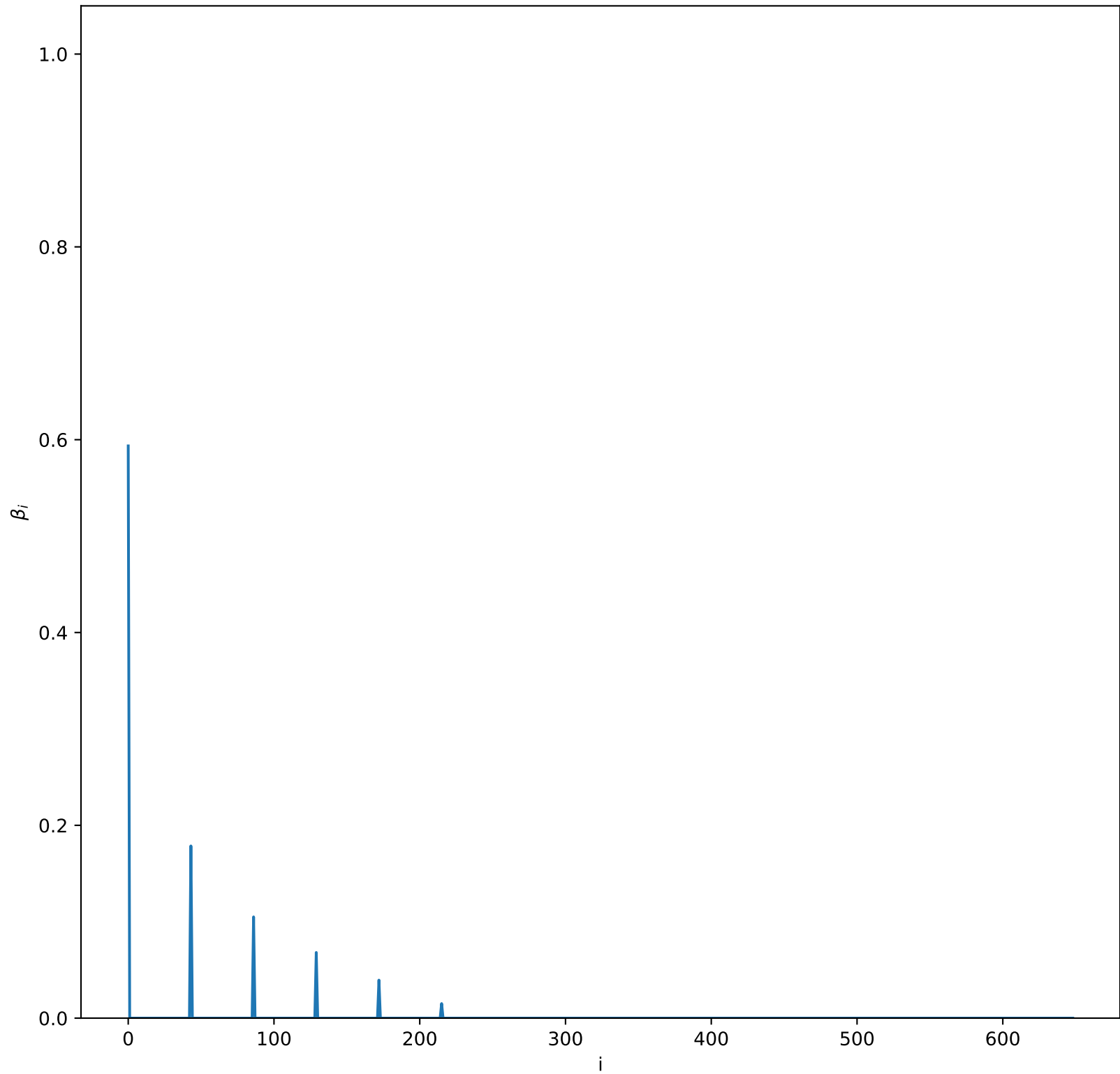
$\mu = 1.25$



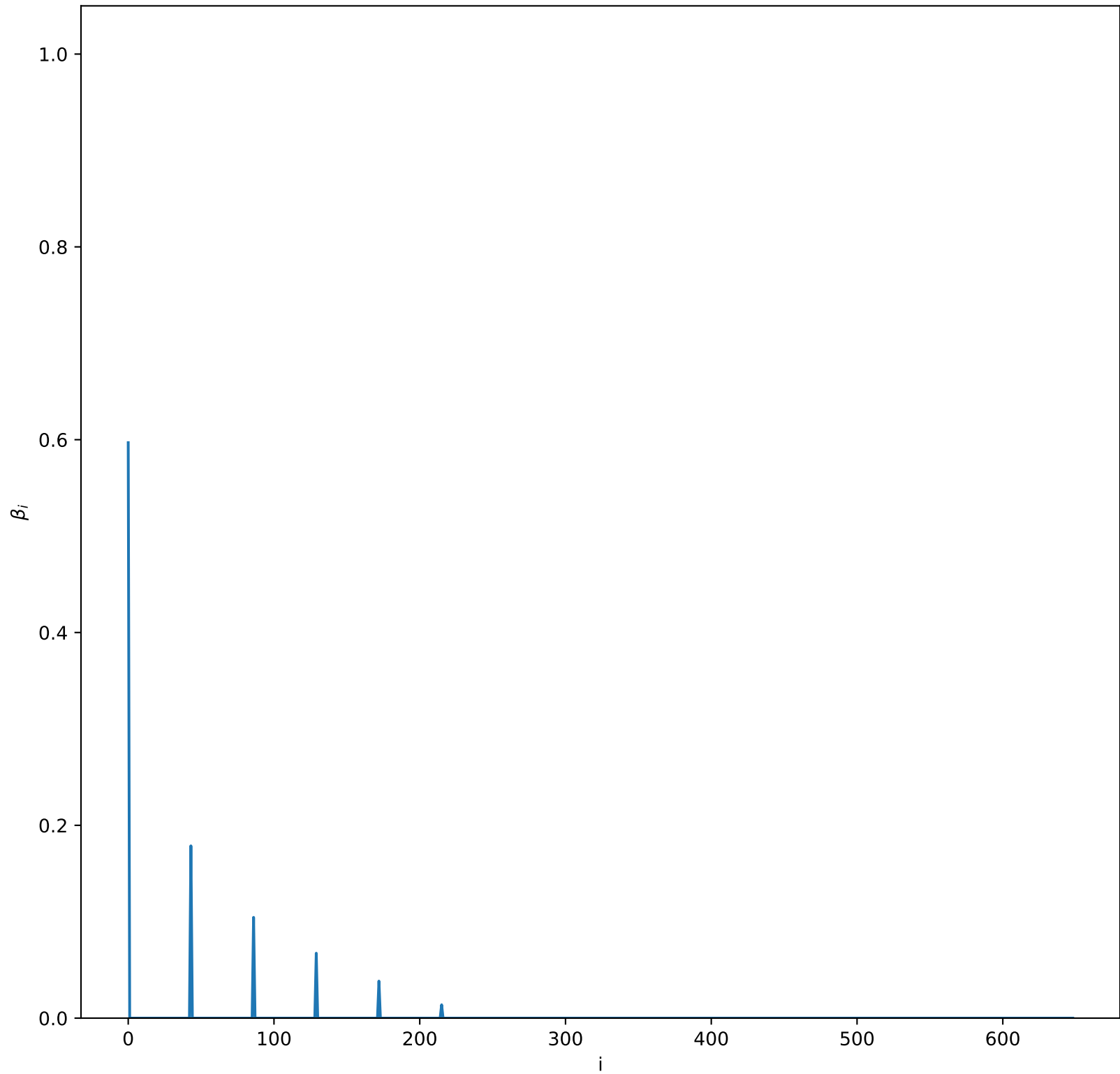
$\mu = 1.26$



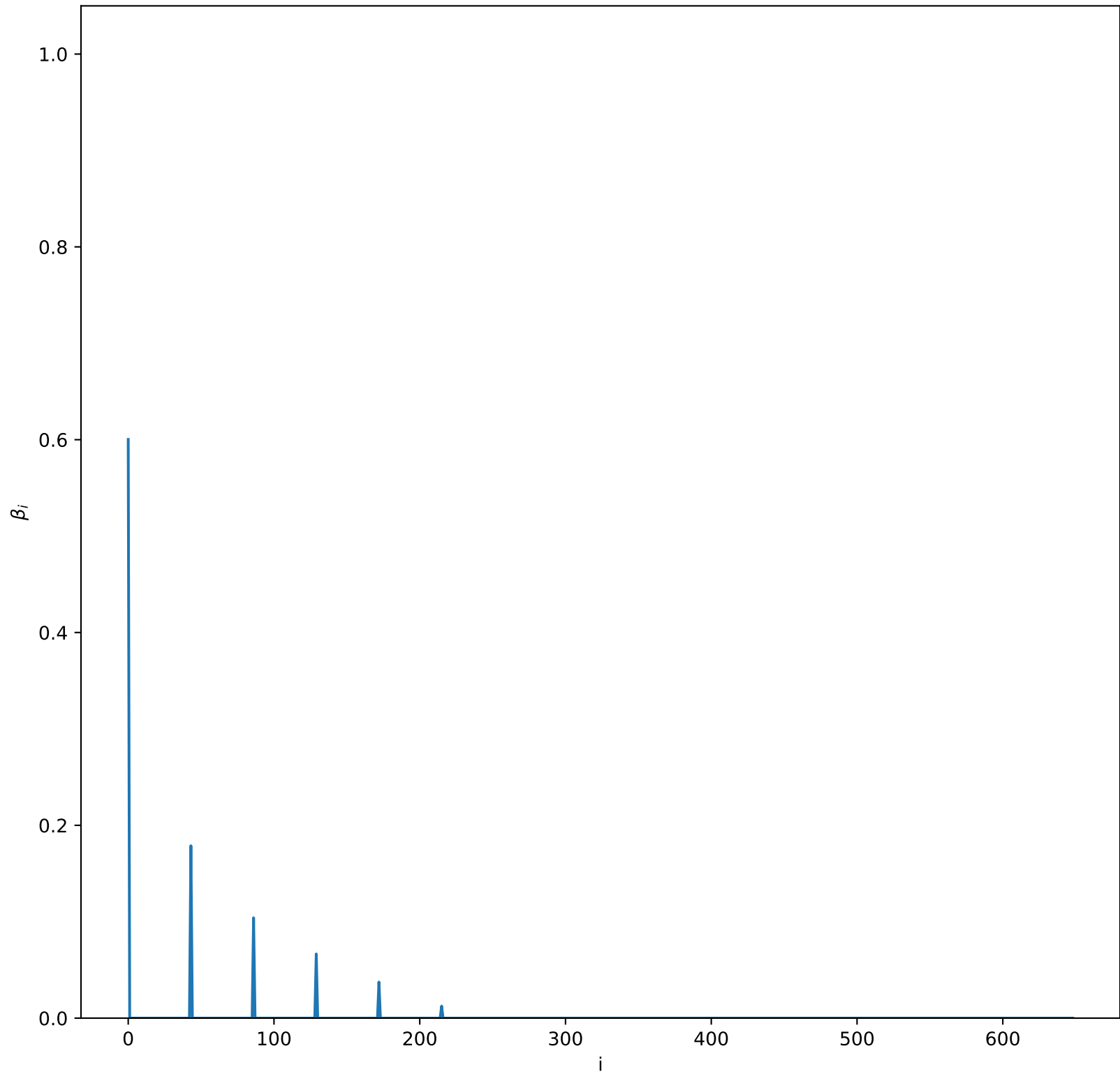
$\mu = 1.27$



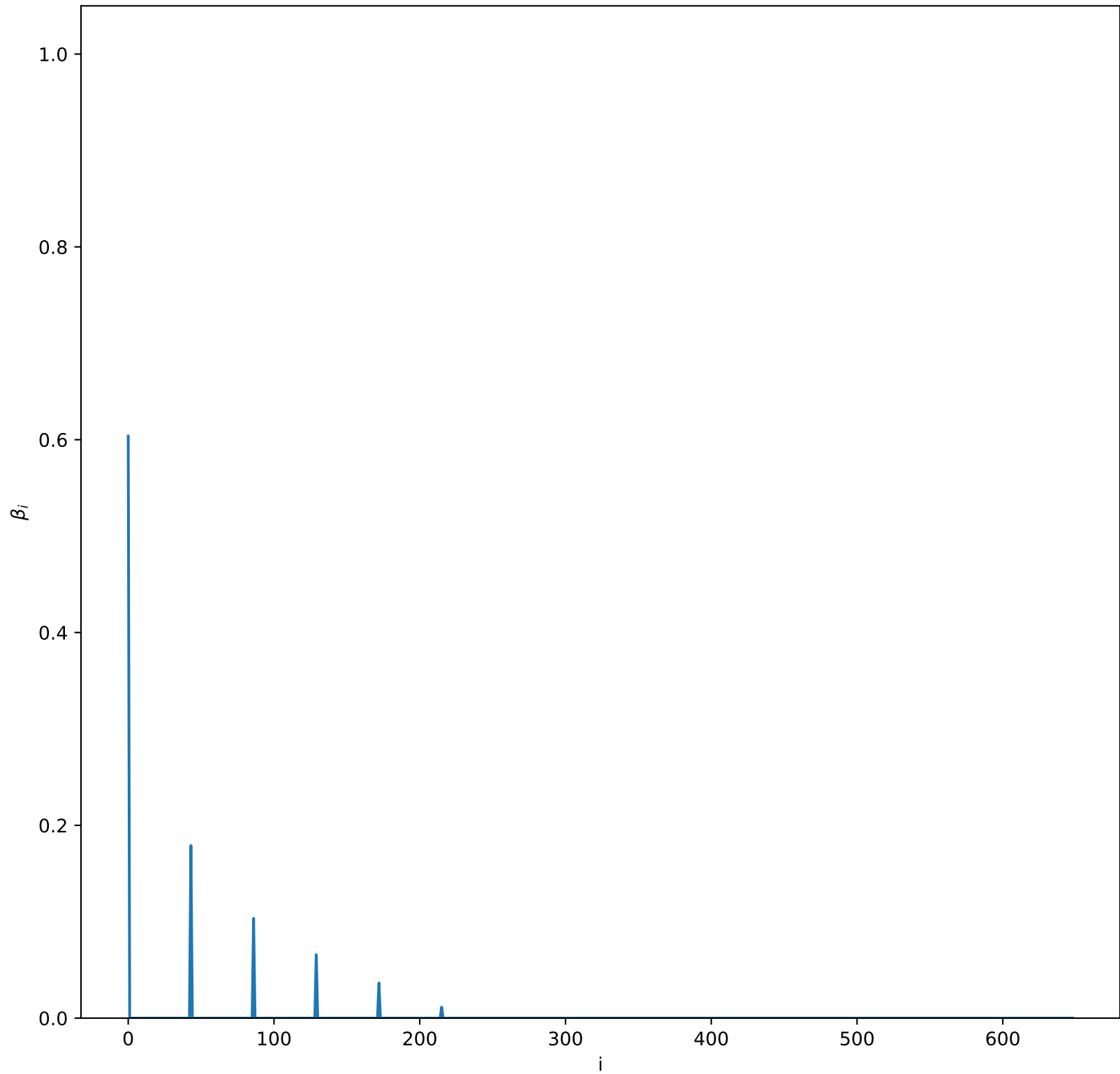
$\mu = 1.28$



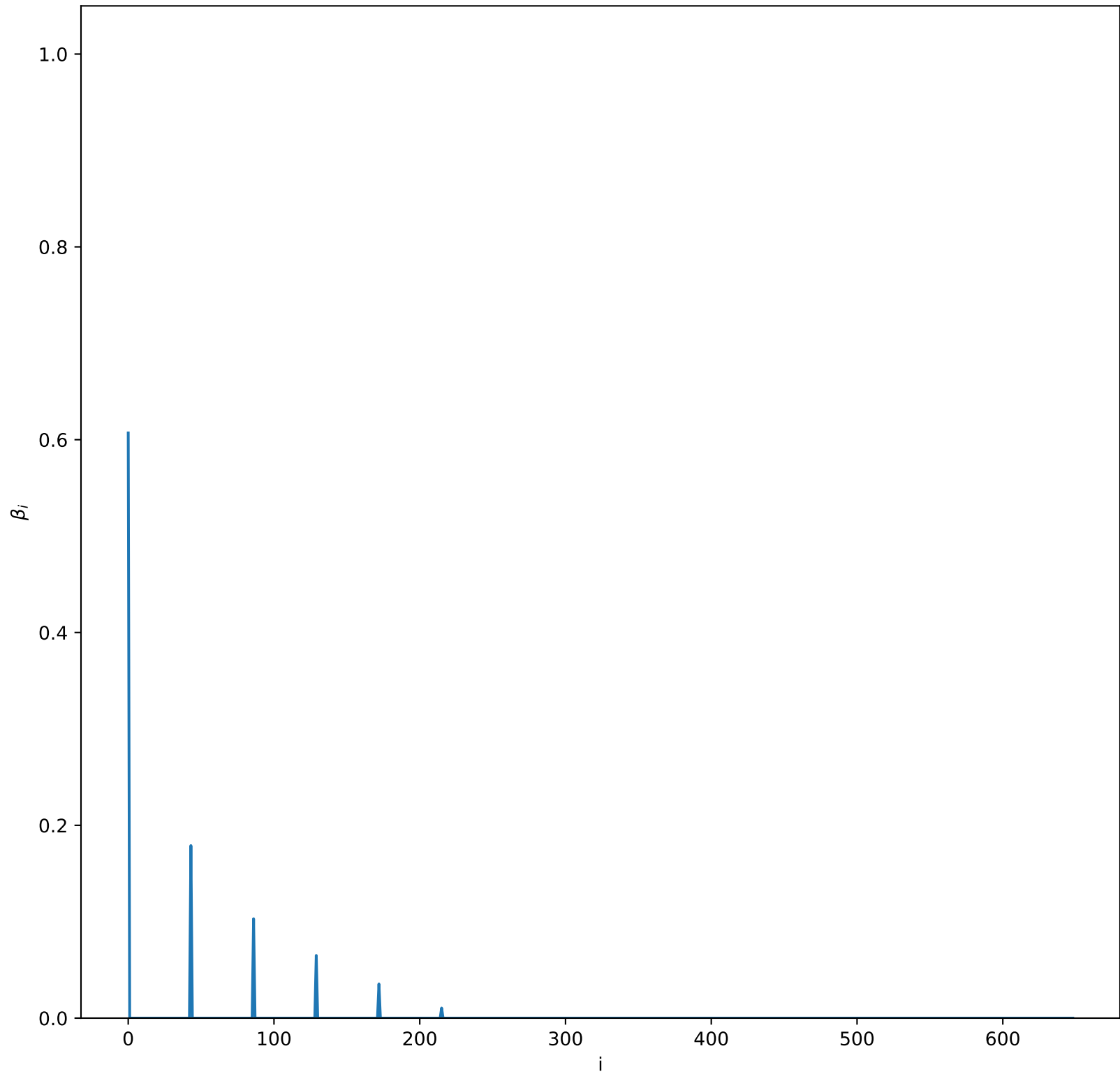
$\mu = 1.29$



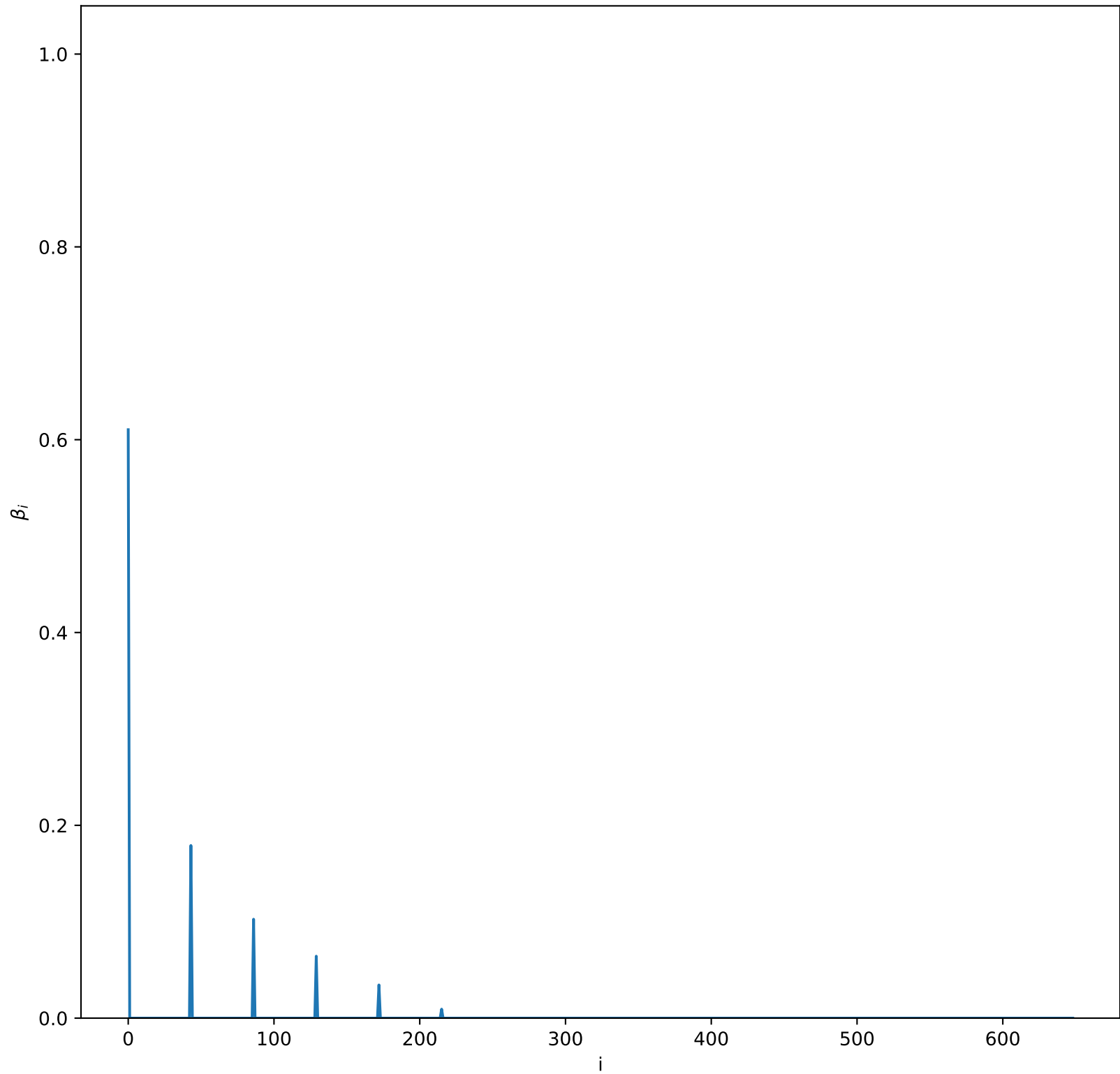
$\mu = 1.30$



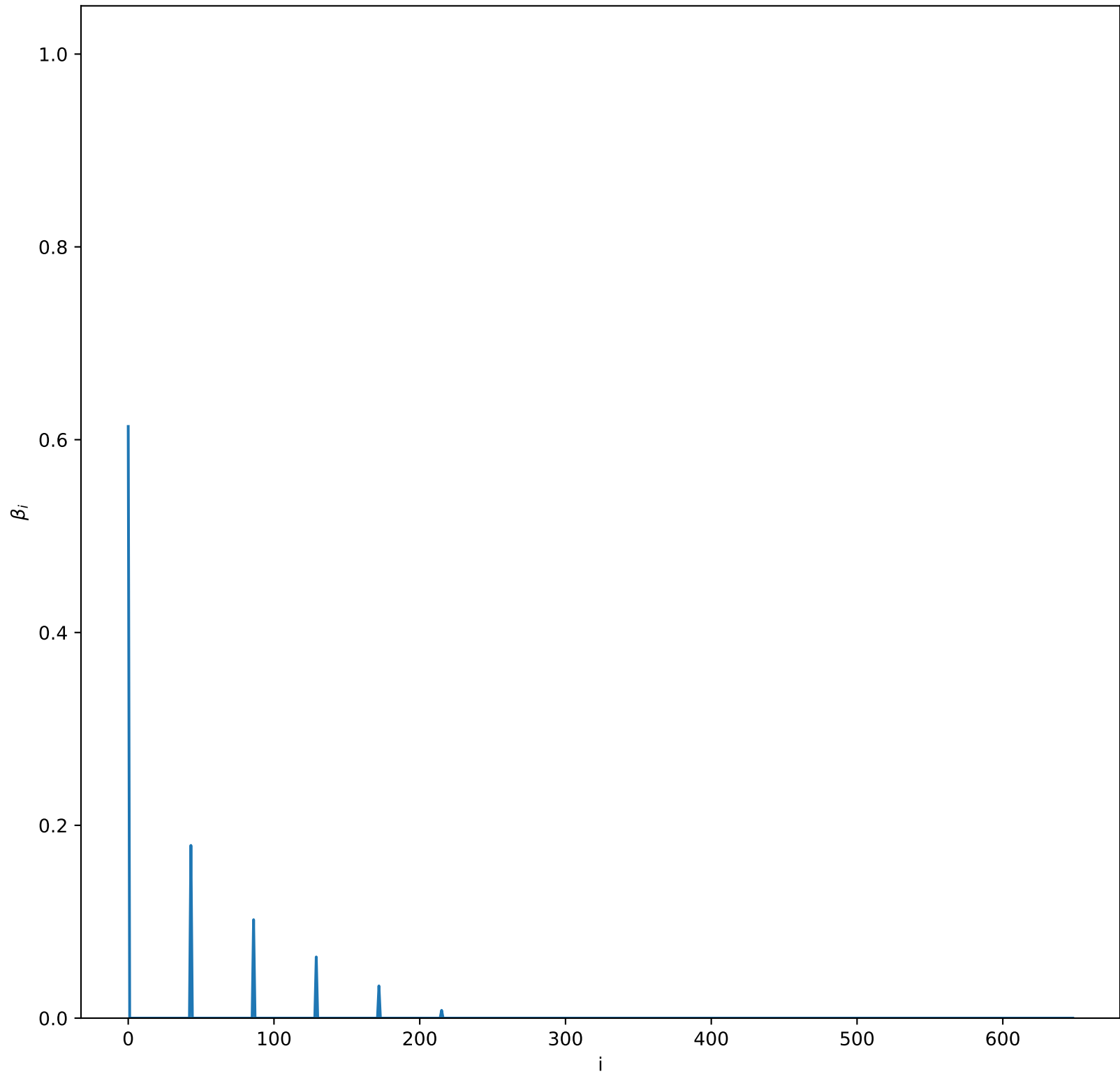
$\mu = 1.31$



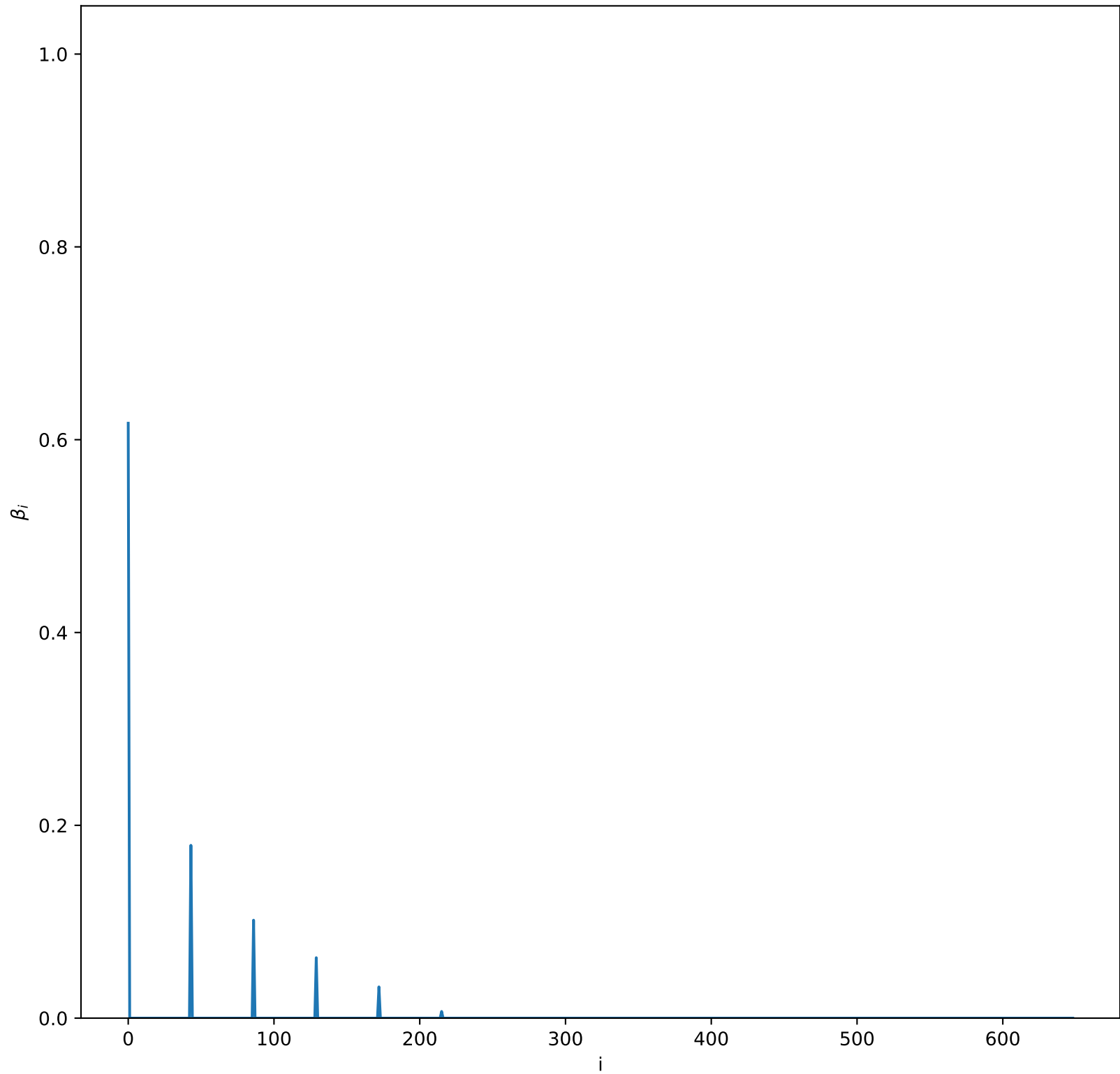
$\mu = 1.32$



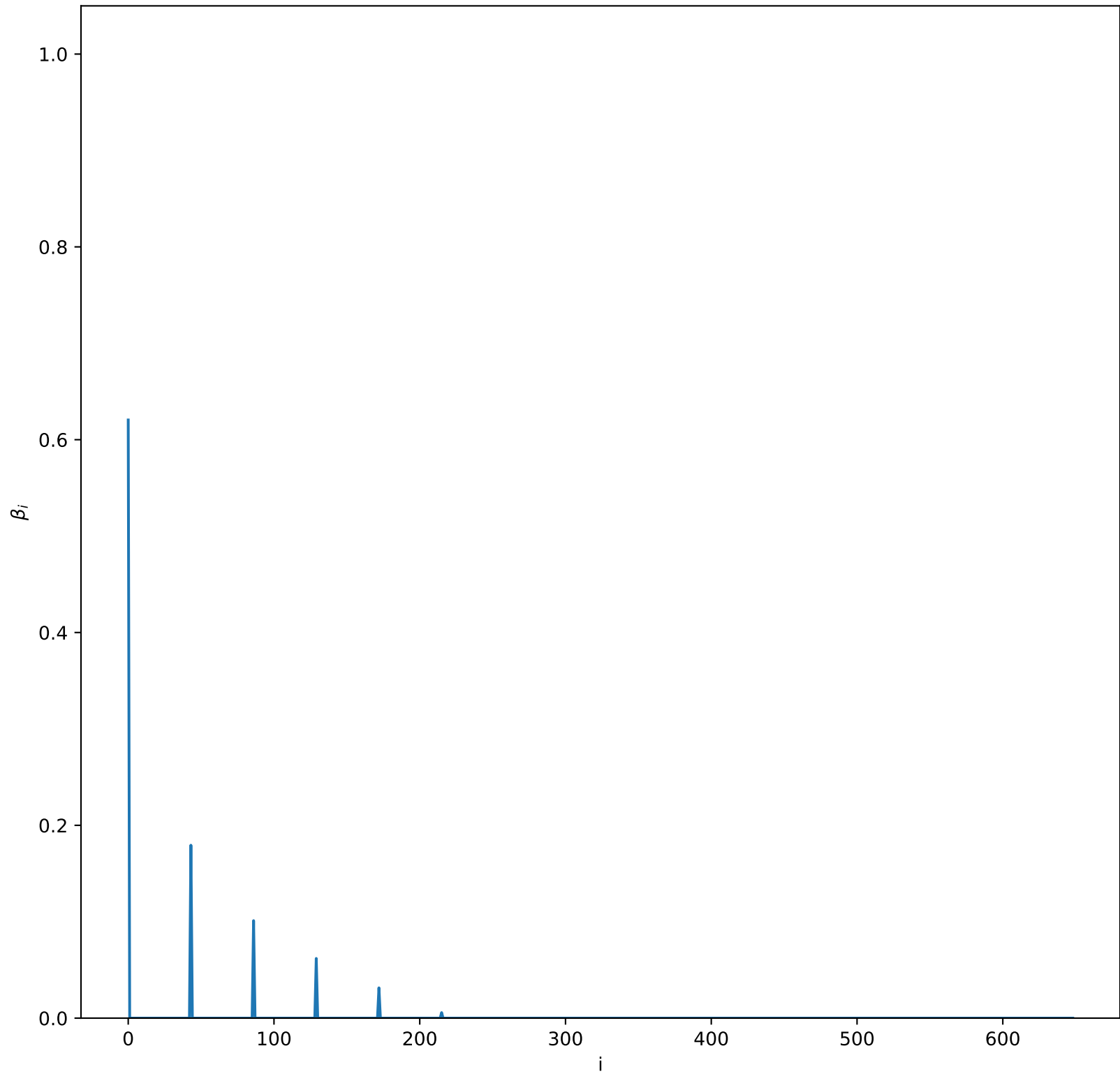
$\mu = 1.33$



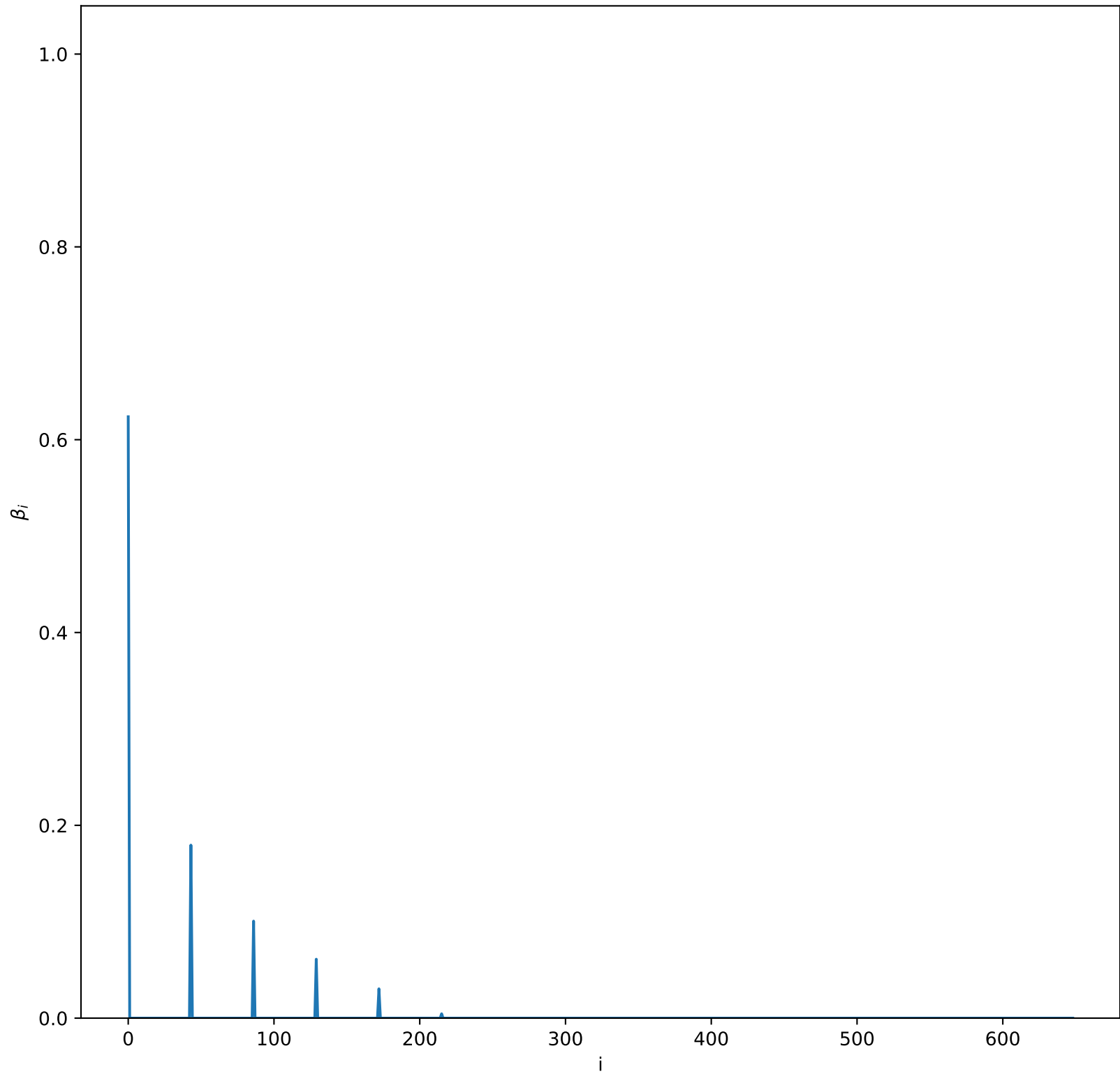
$\mu = 1.34$



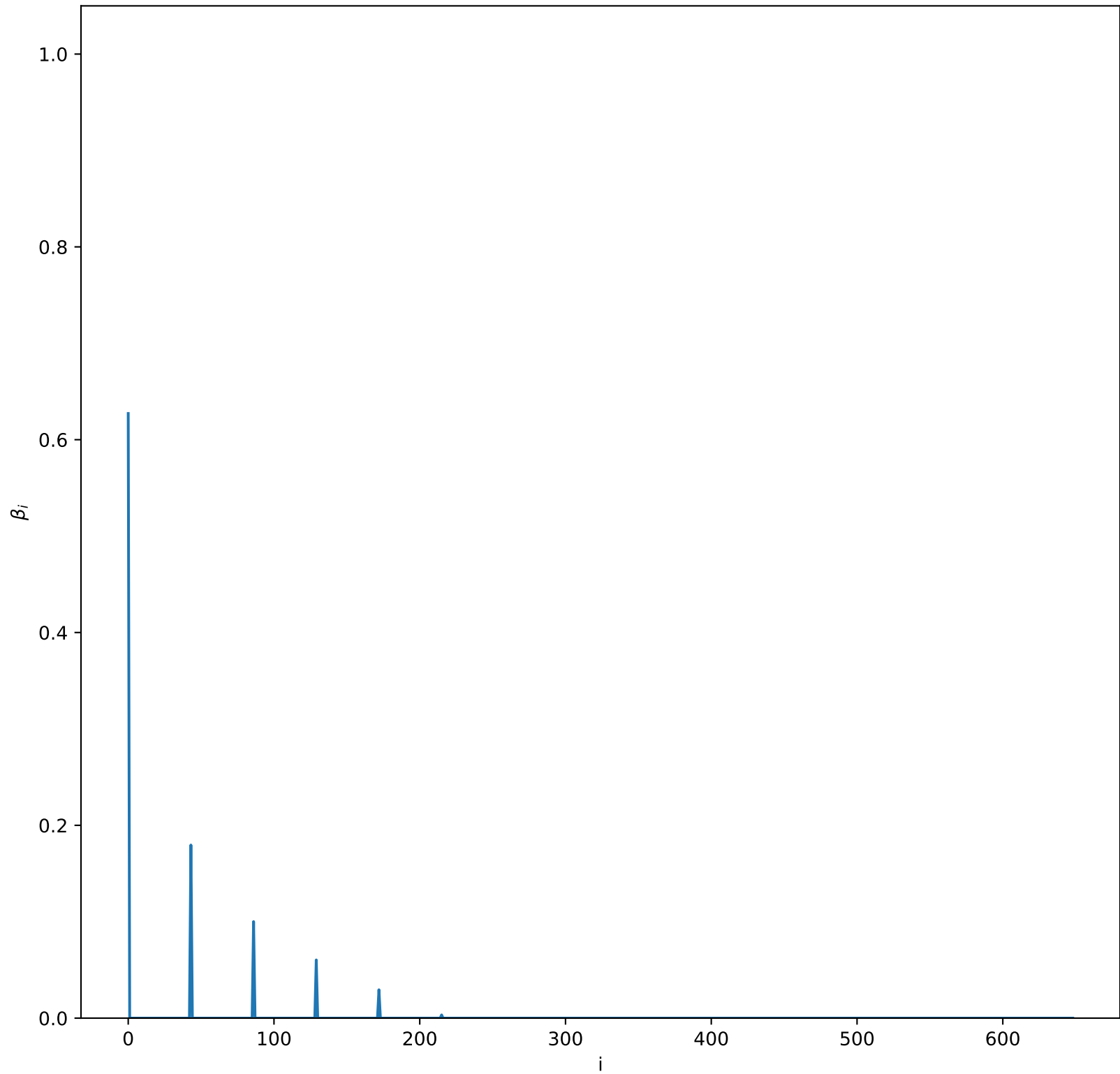
$\mu = 1.35$



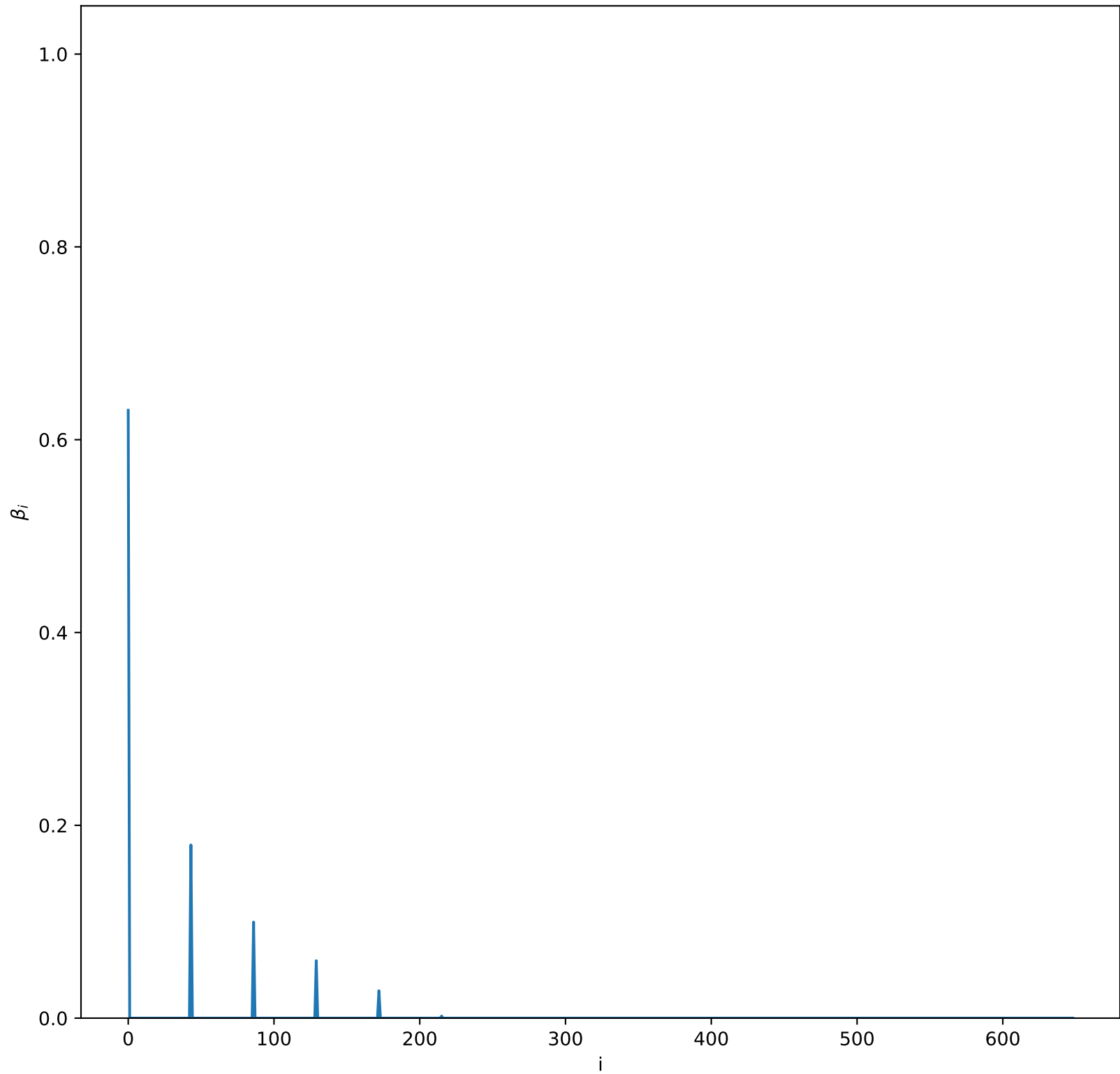
$\mu = 1.36$



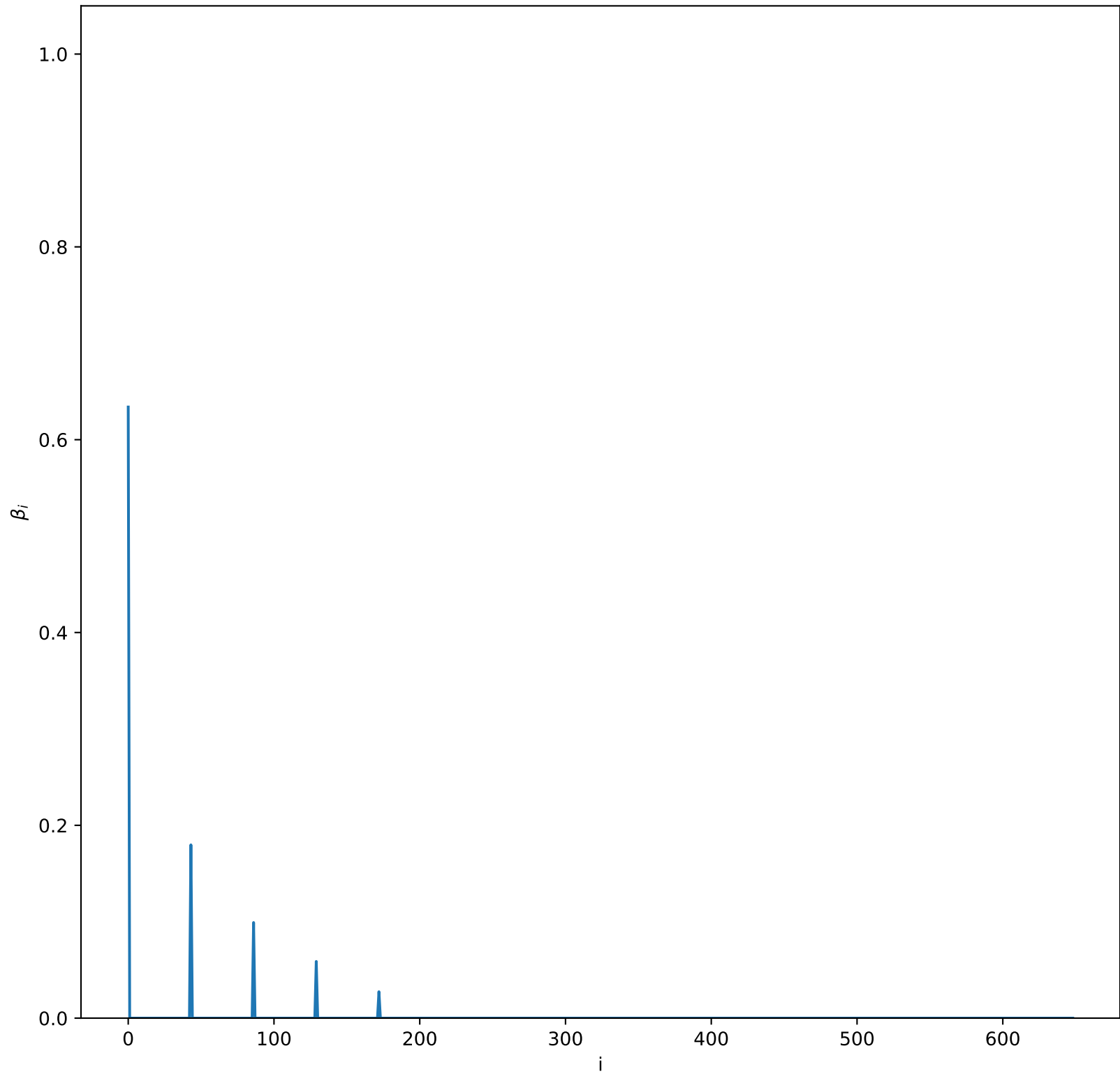
$\mu = 1.37$



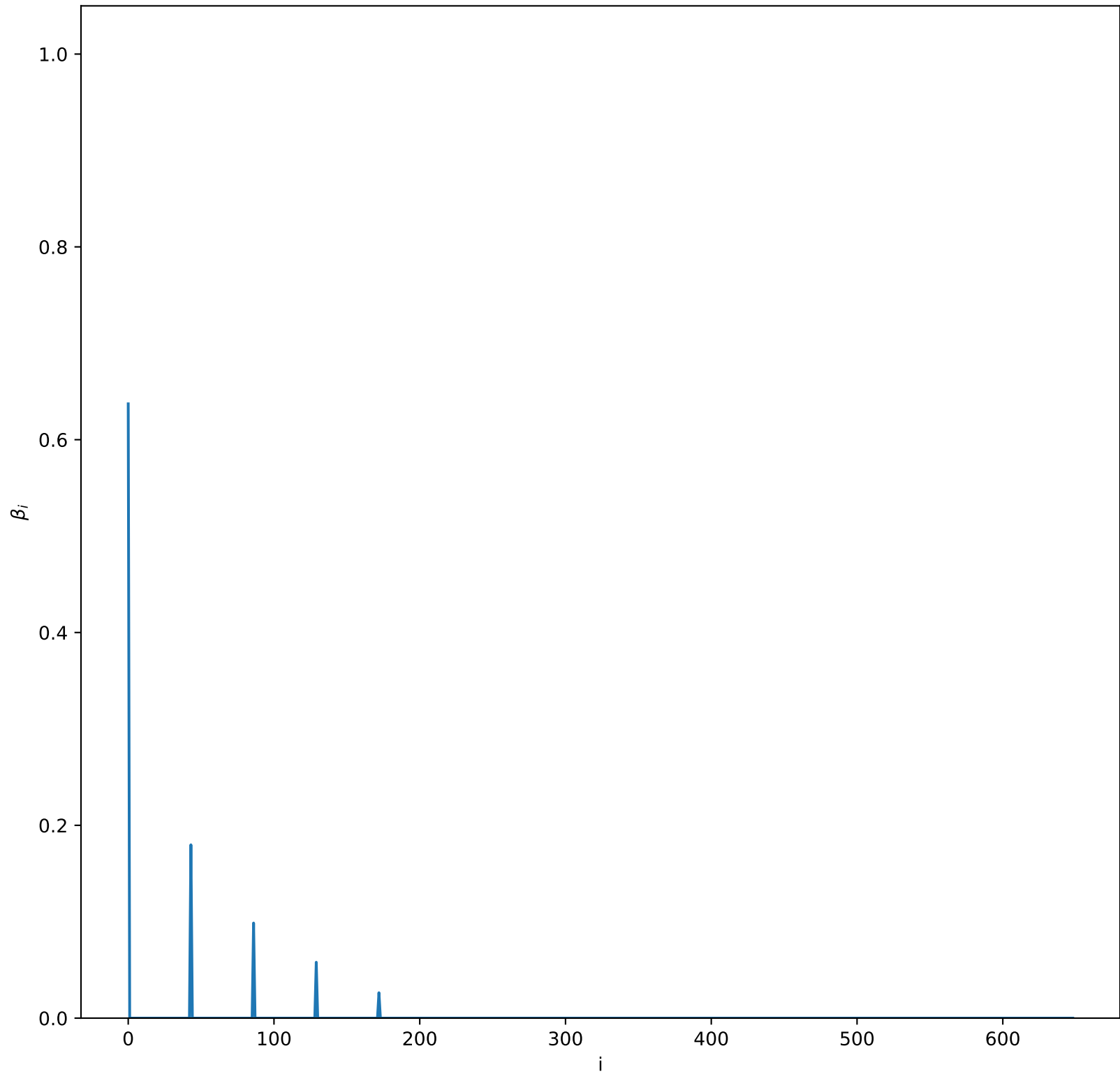
$\mu = 1.38$



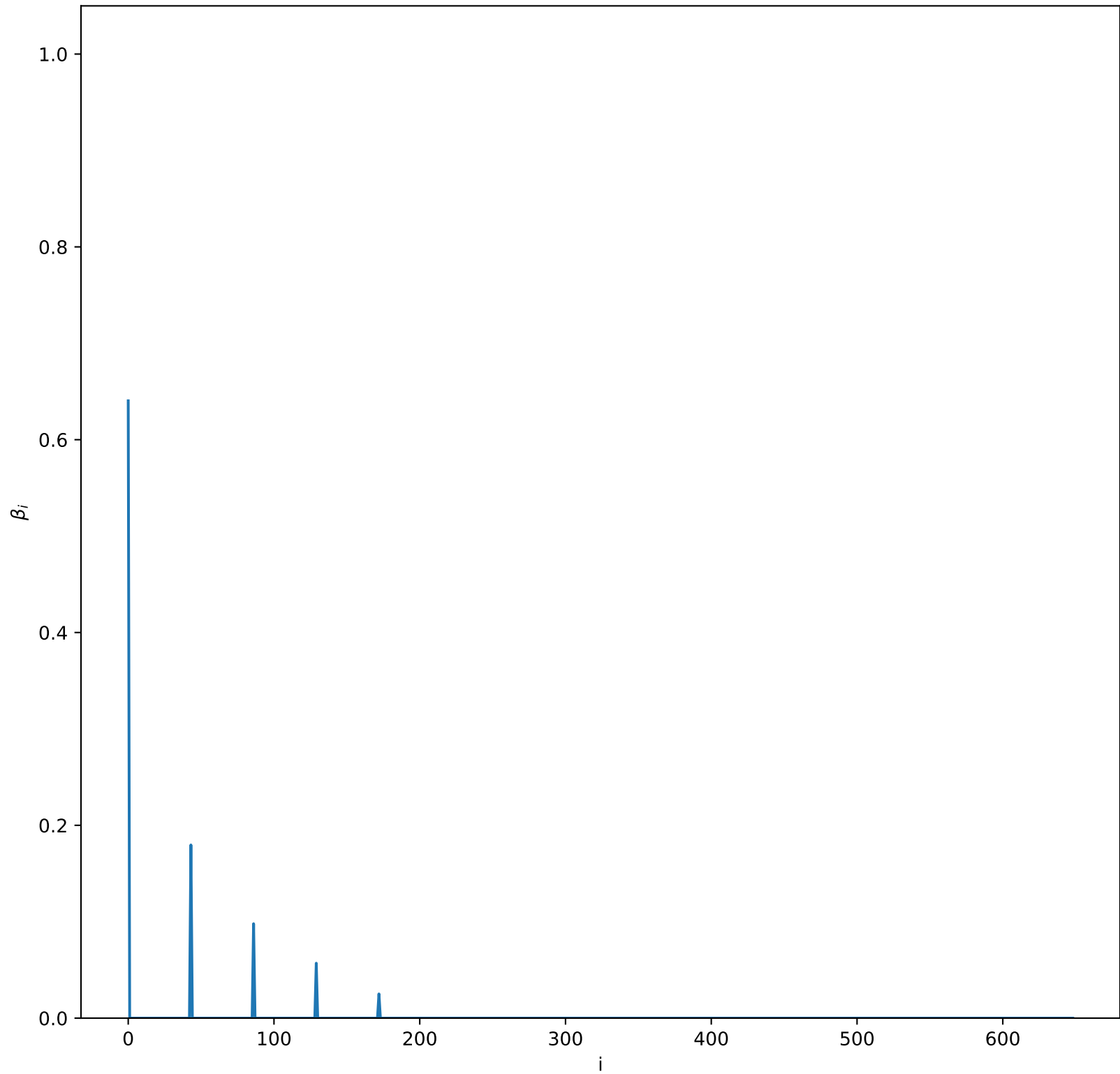
$\mu = 1.39$



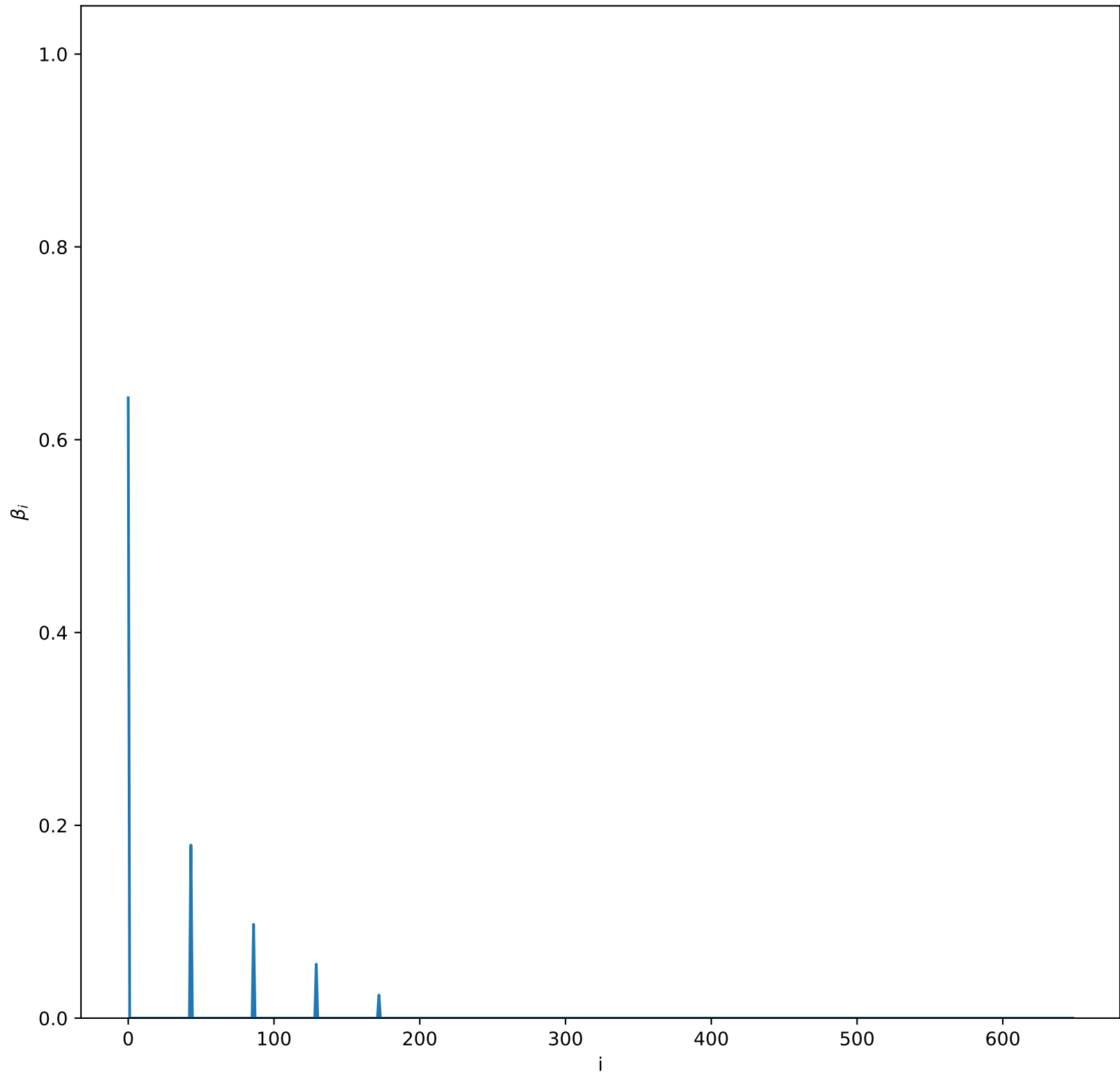
$\mu = 1.40$



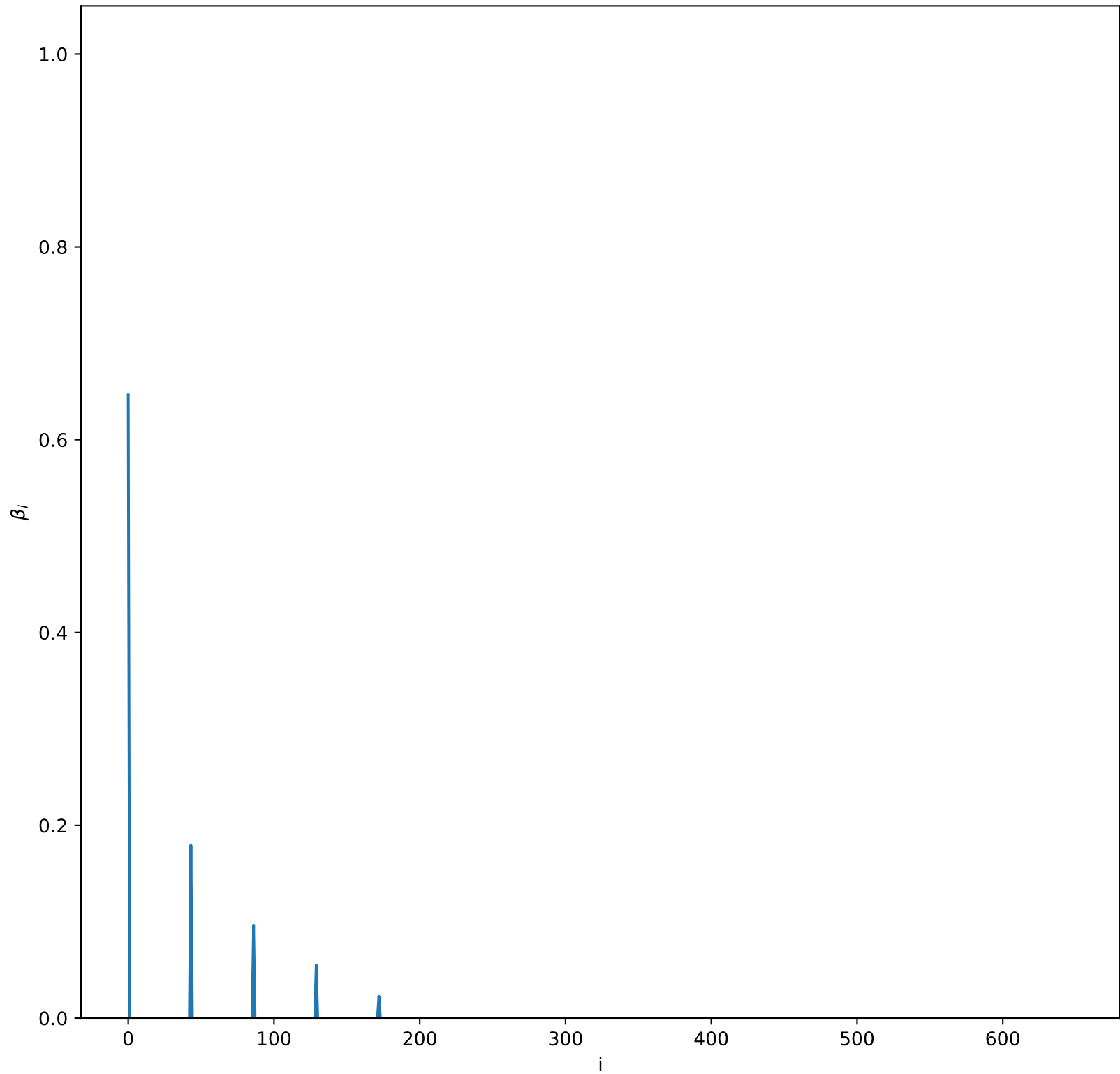
$\mu = 1.41$



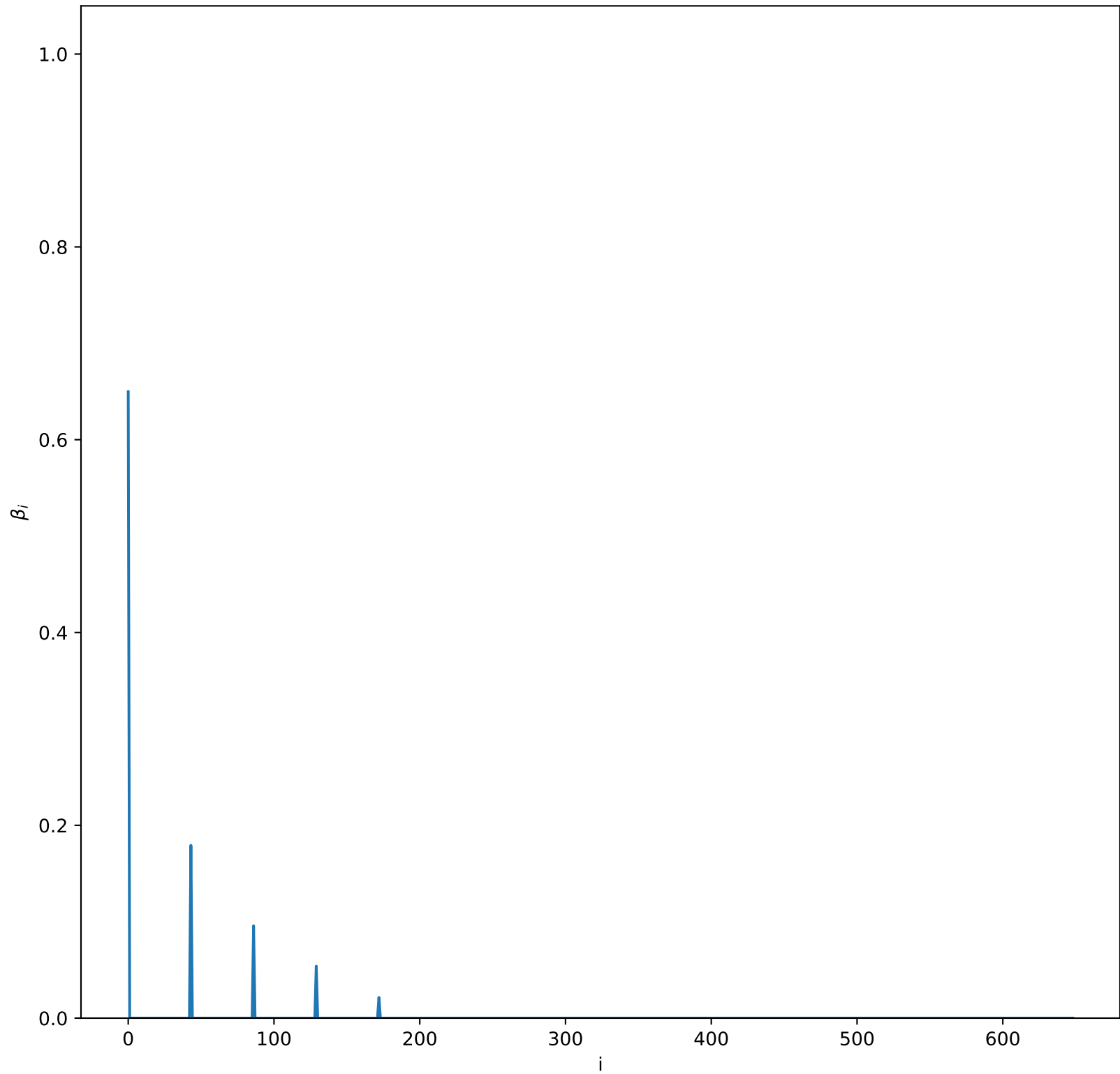
$\mu = 1.42$



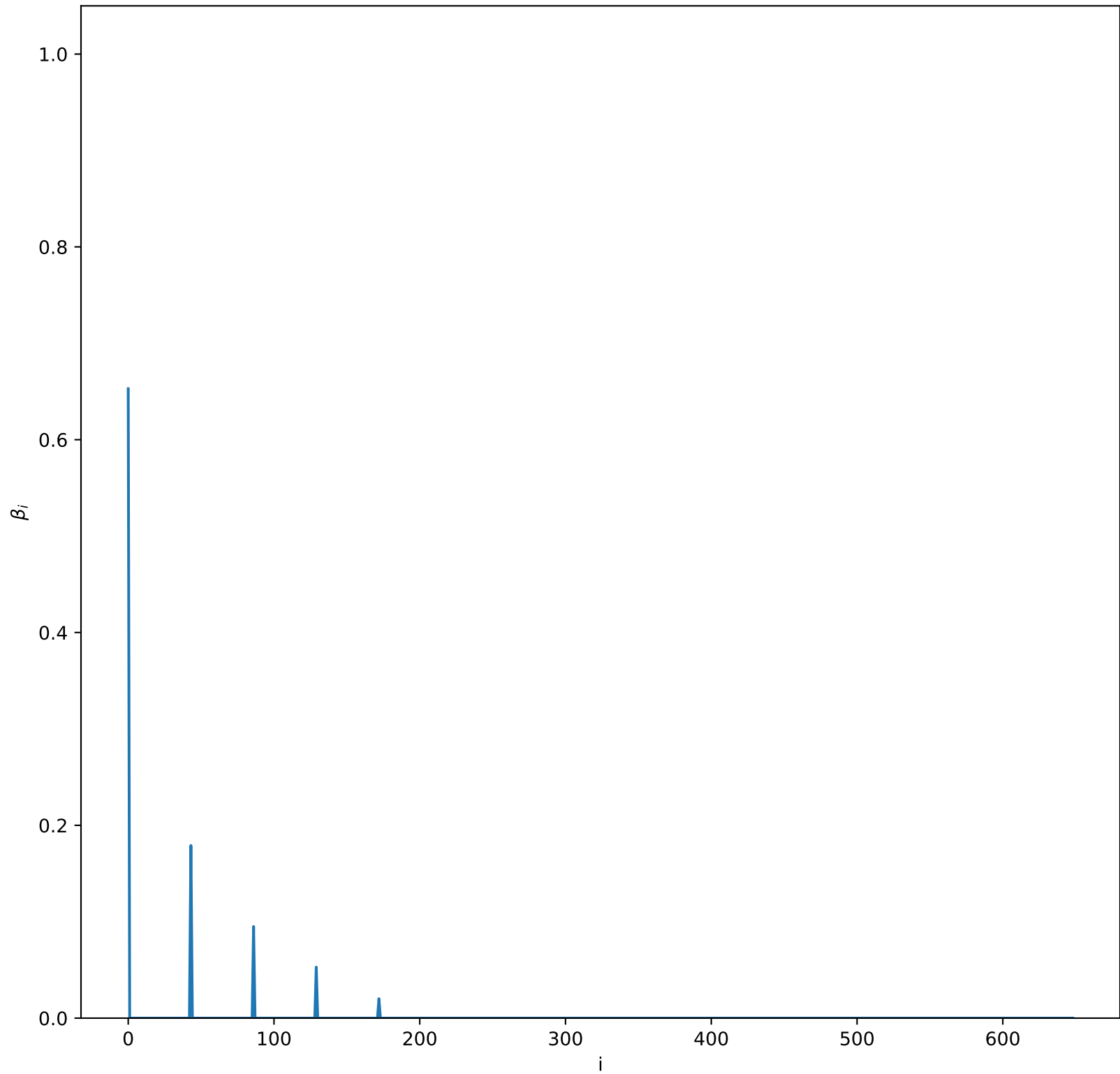
$\mu = 1.43$



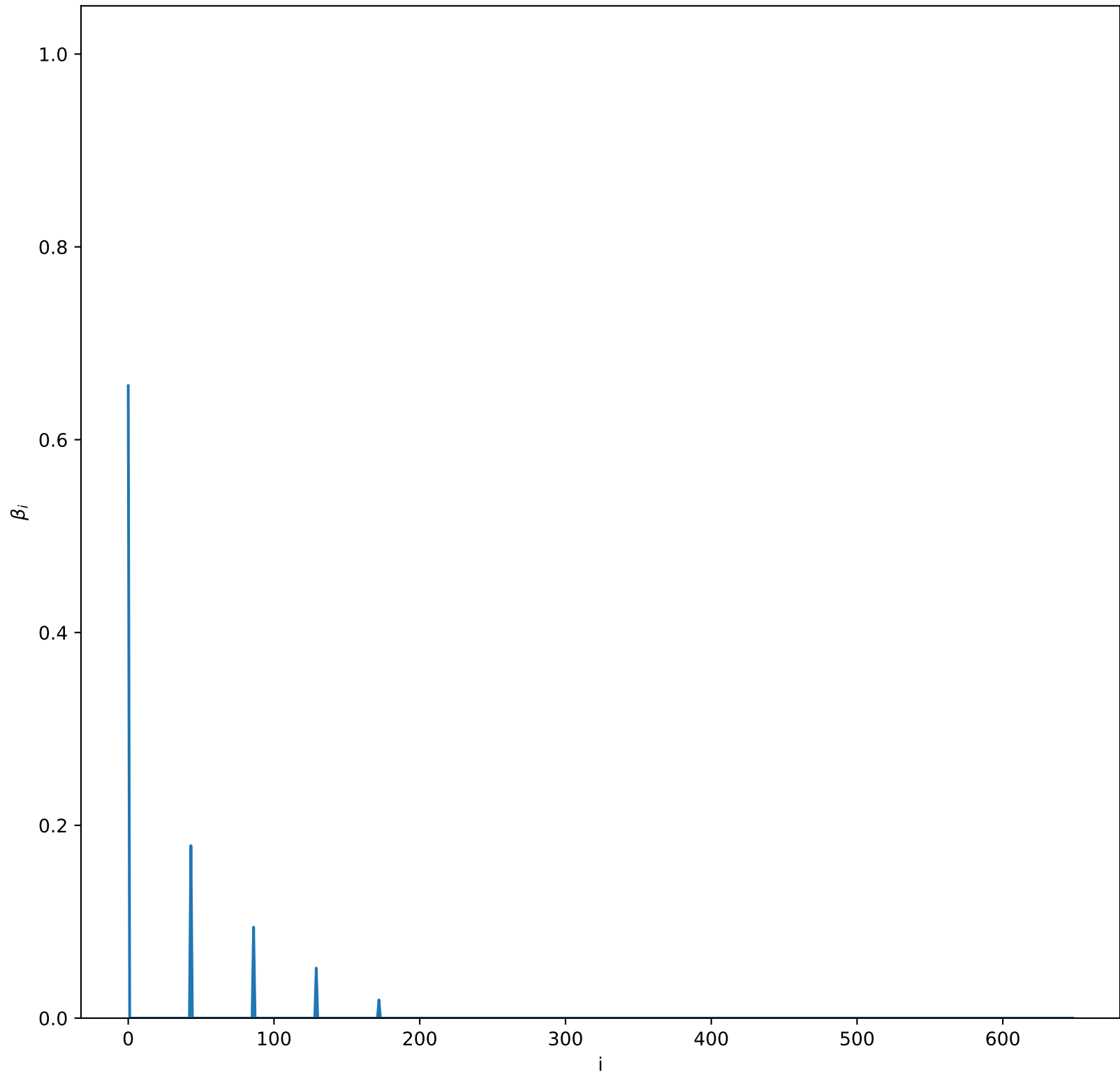
$\mu = 1.44$



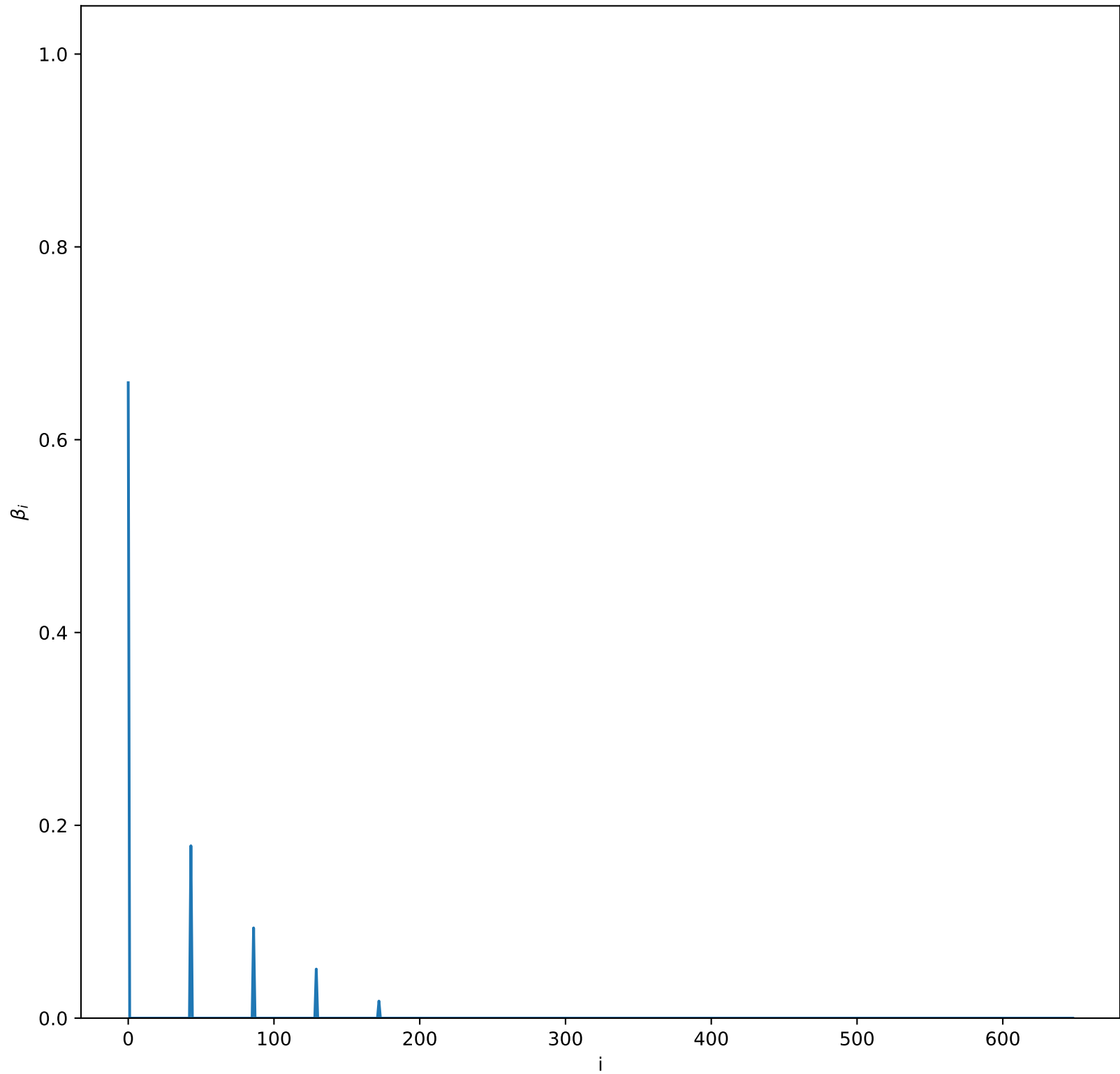
$\mu = 1.45$



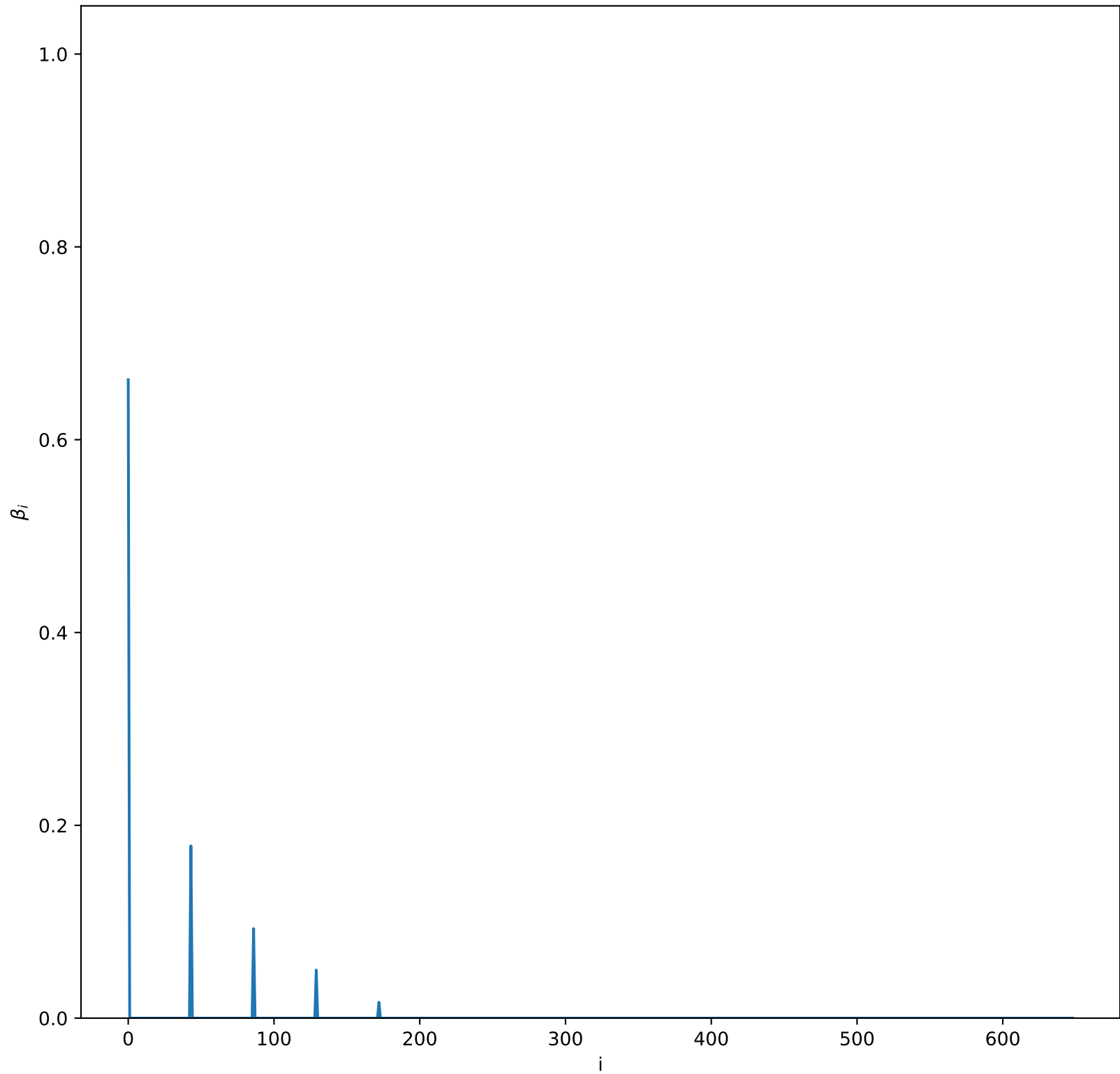
$\mu = 1.46$



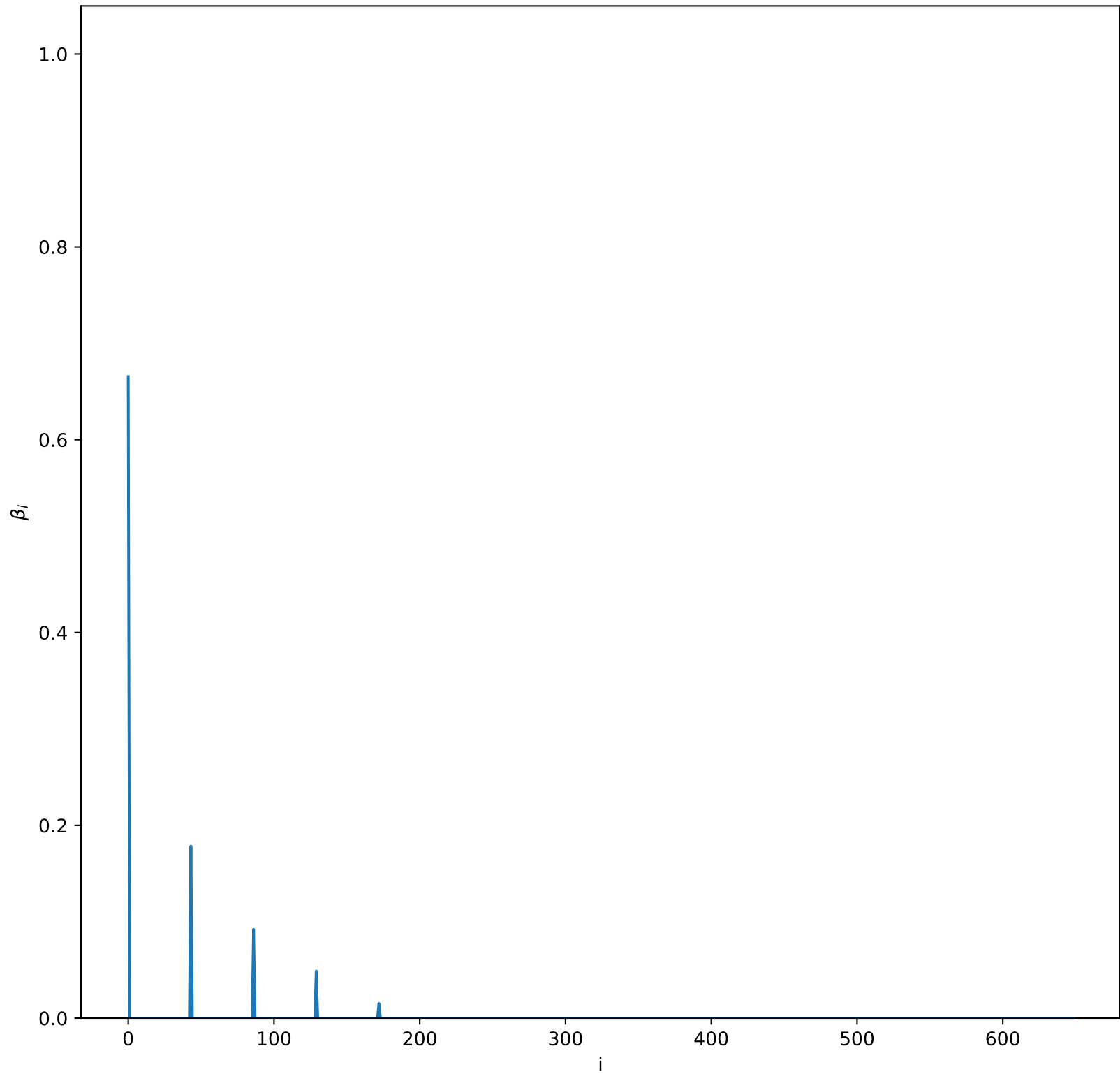
$\mu = 1.47$



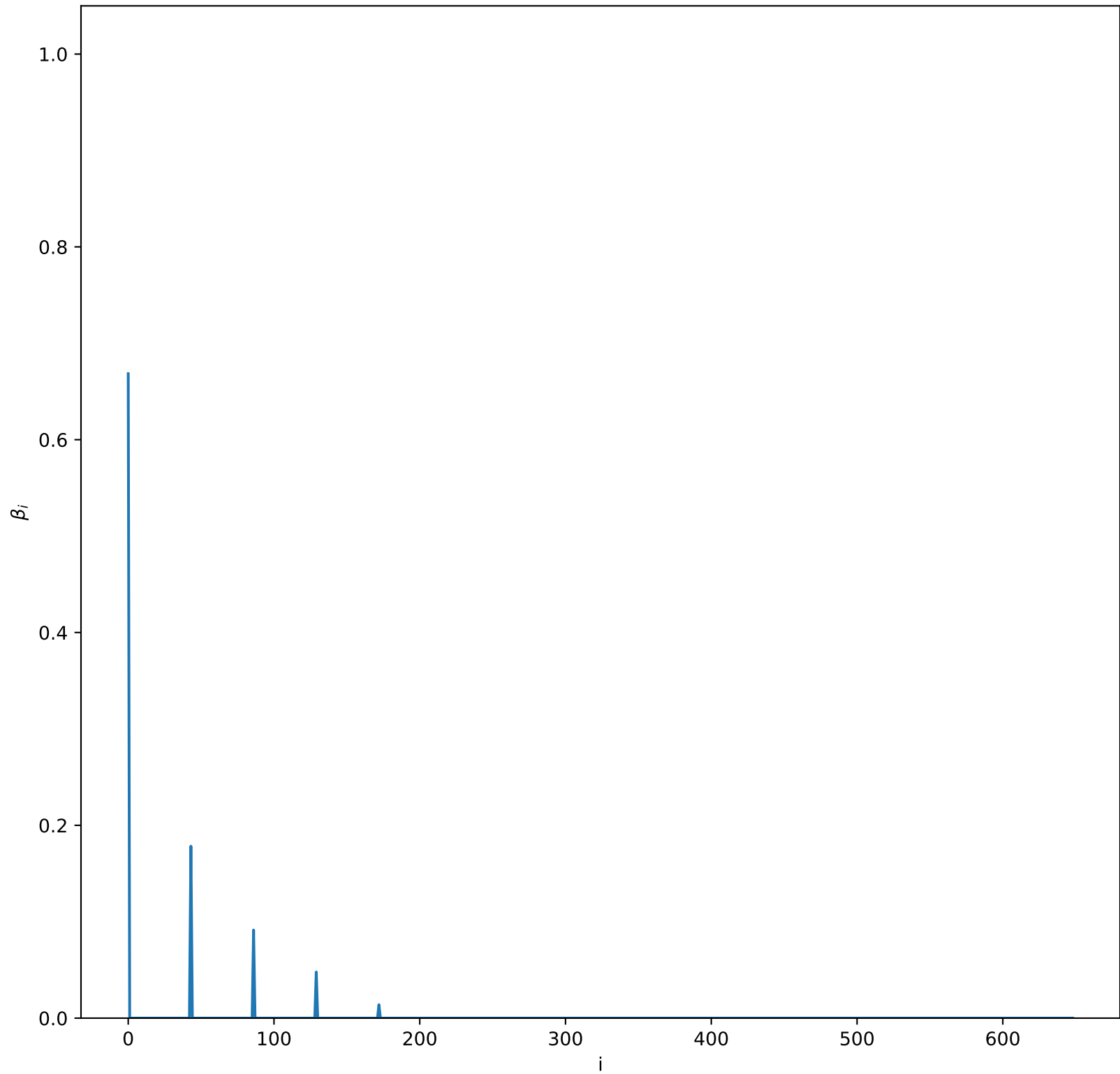
$\mu = 1.48$



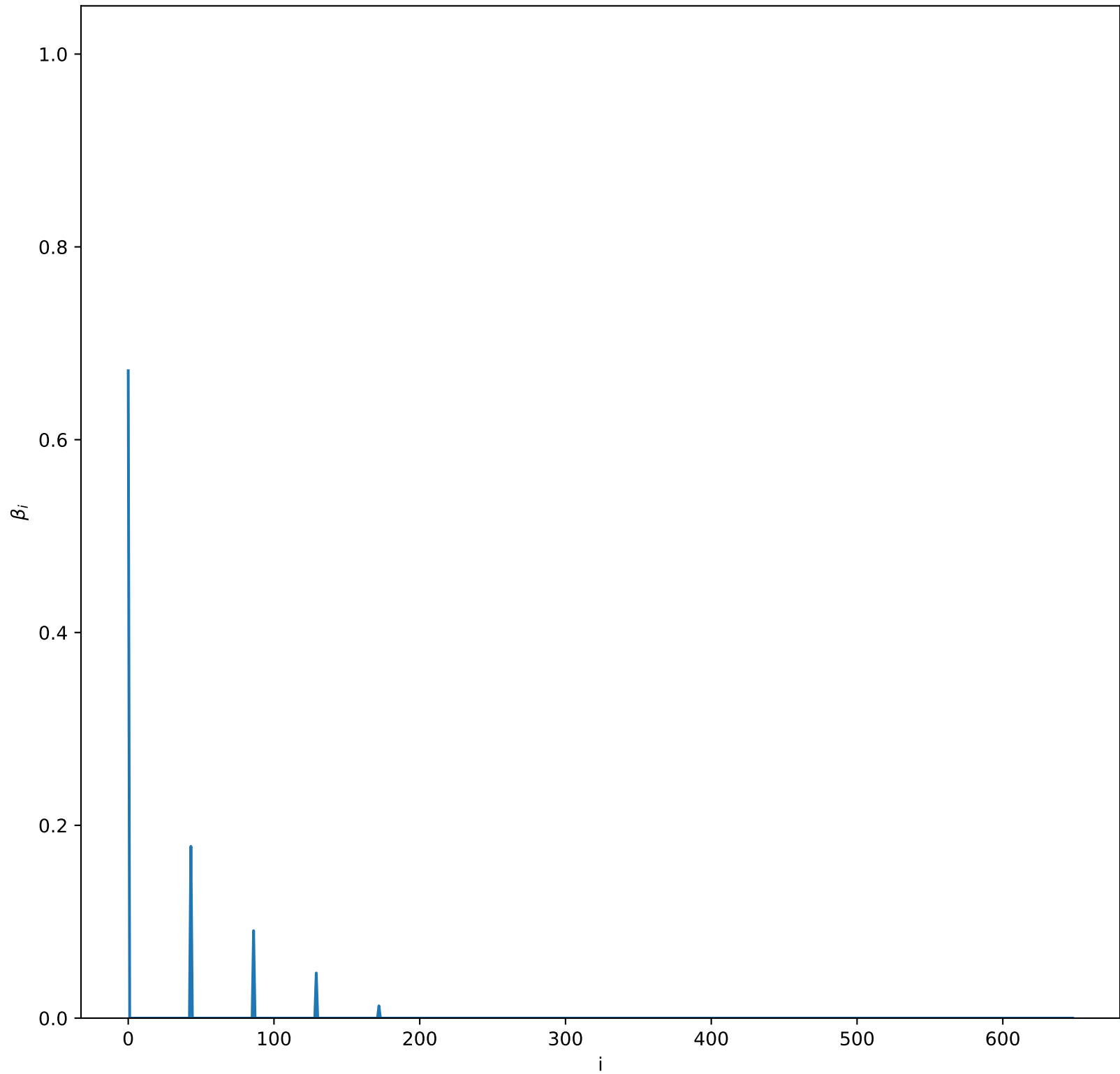
$\mu = 1.49$



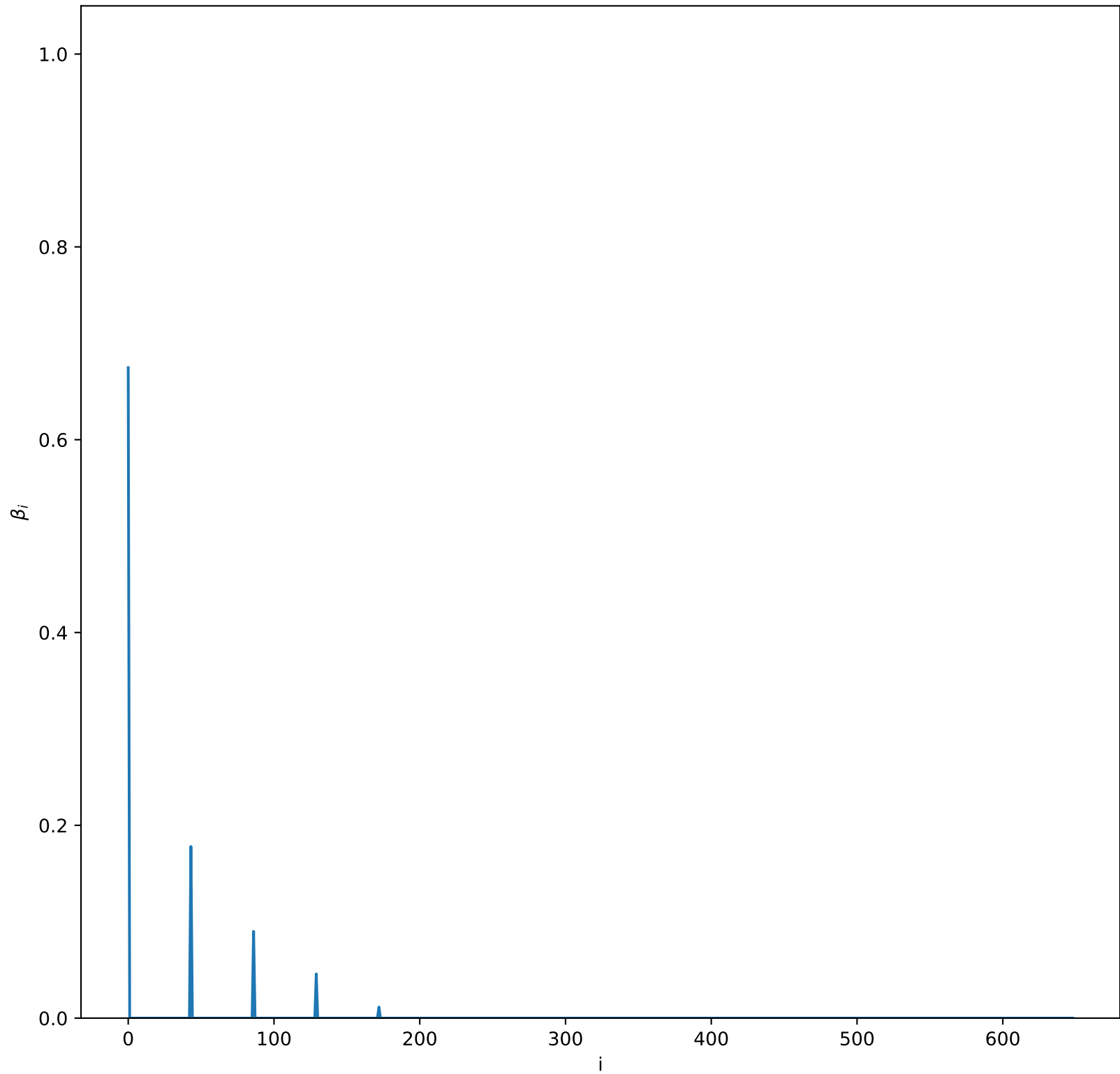
$\mu = 1.50$



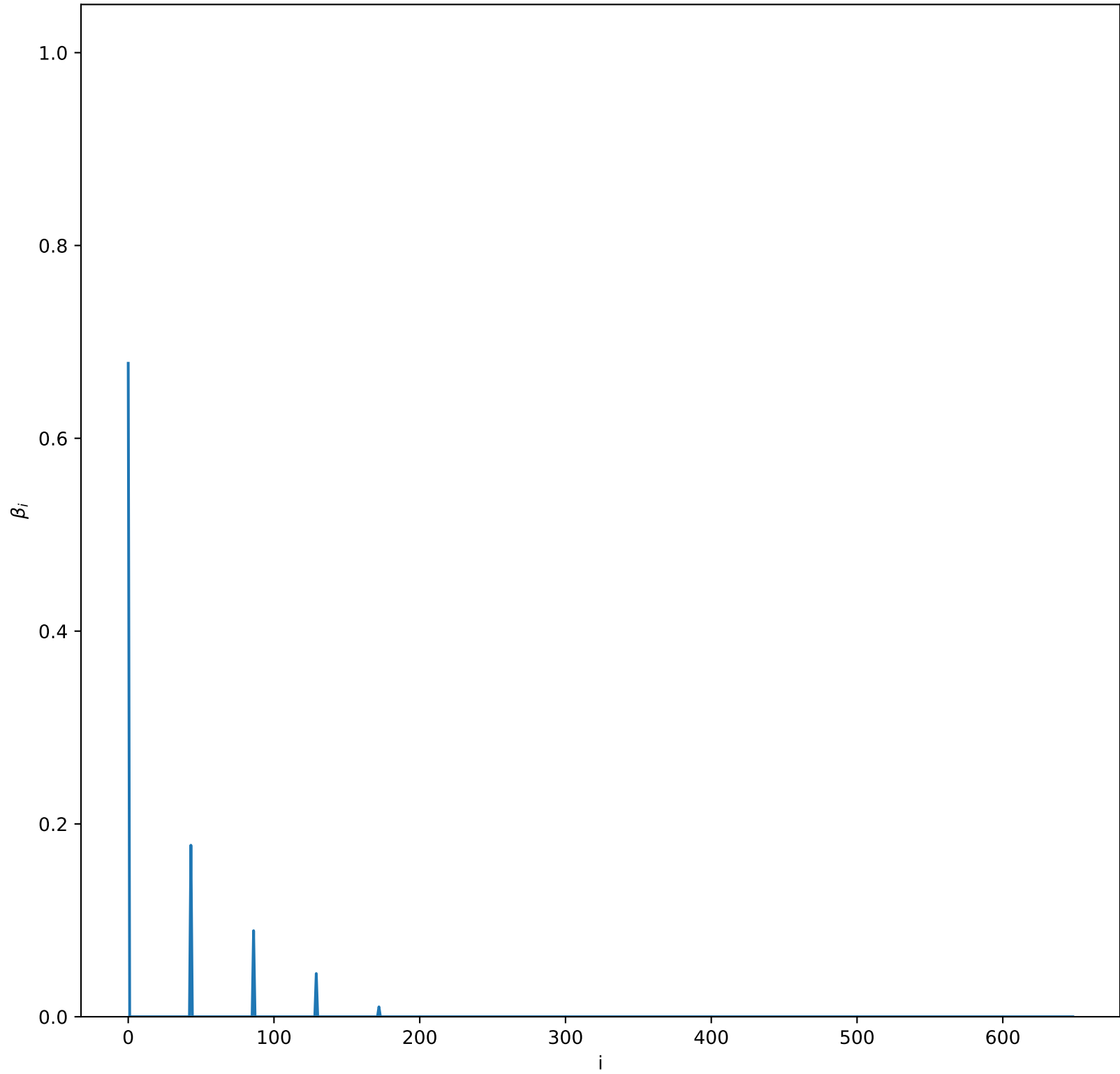
$\mu = 1.51$



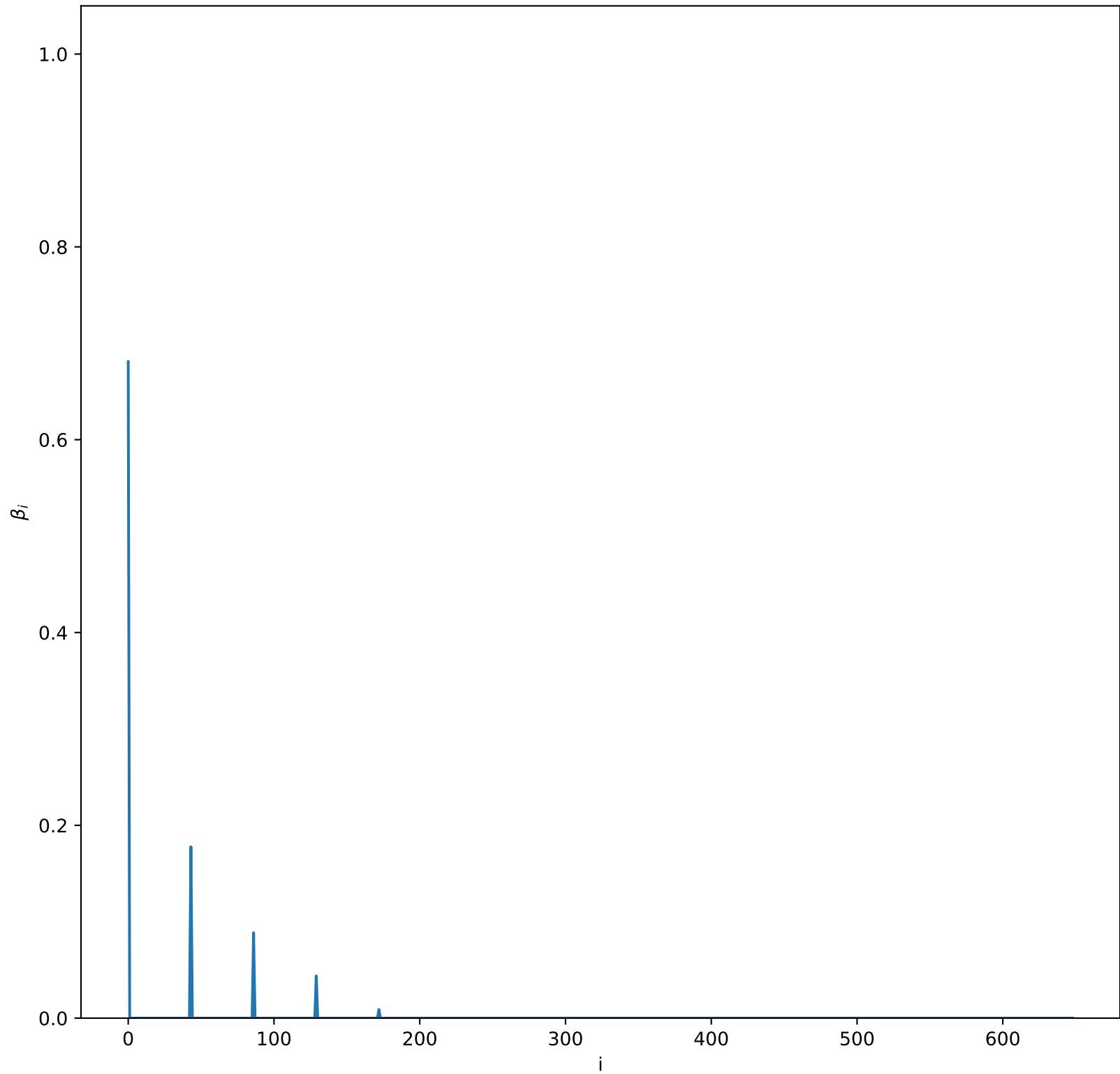
$\mu = 1.52$



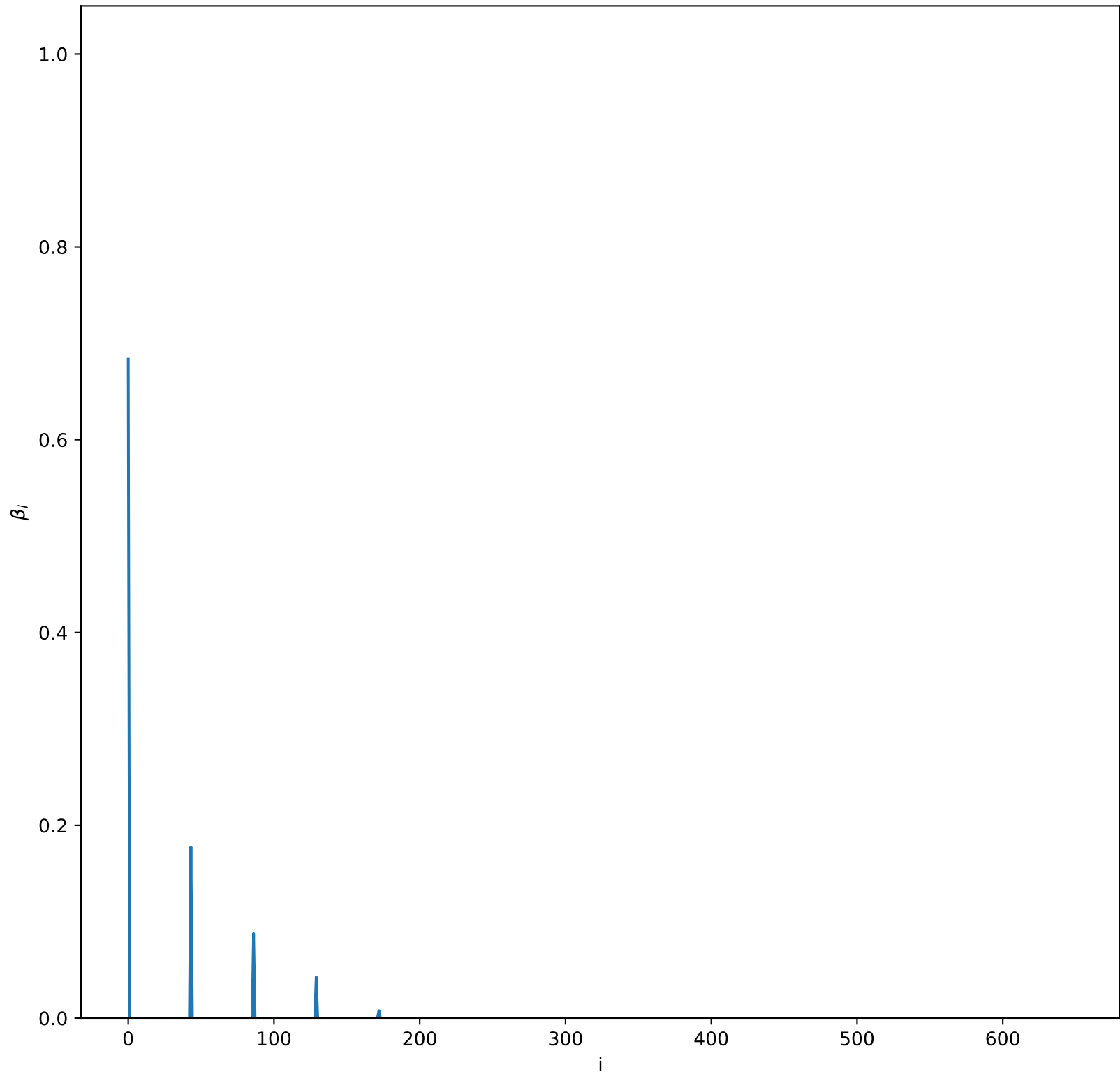
$\mu = 1.53$



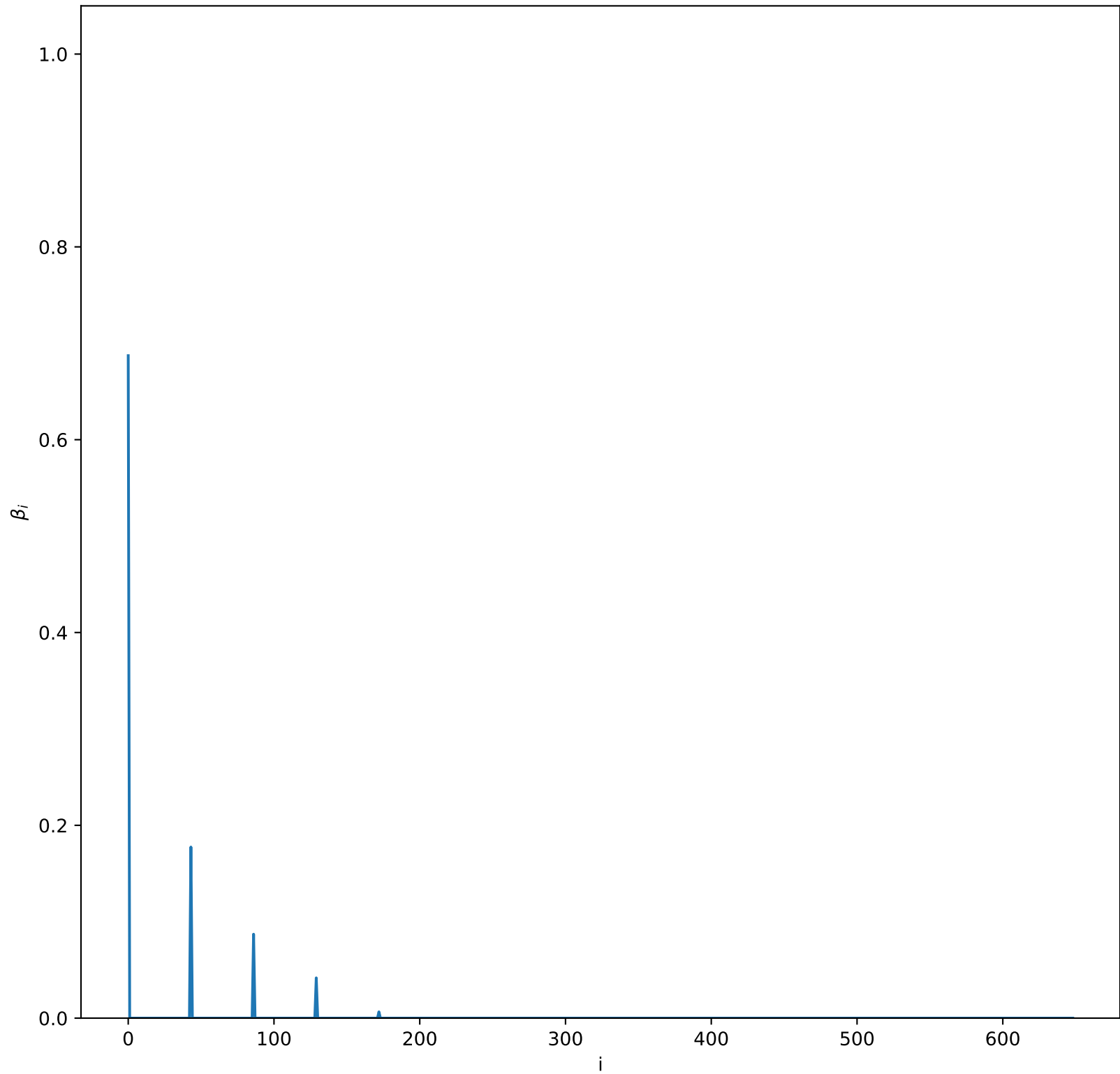
$\mu = 1.54$



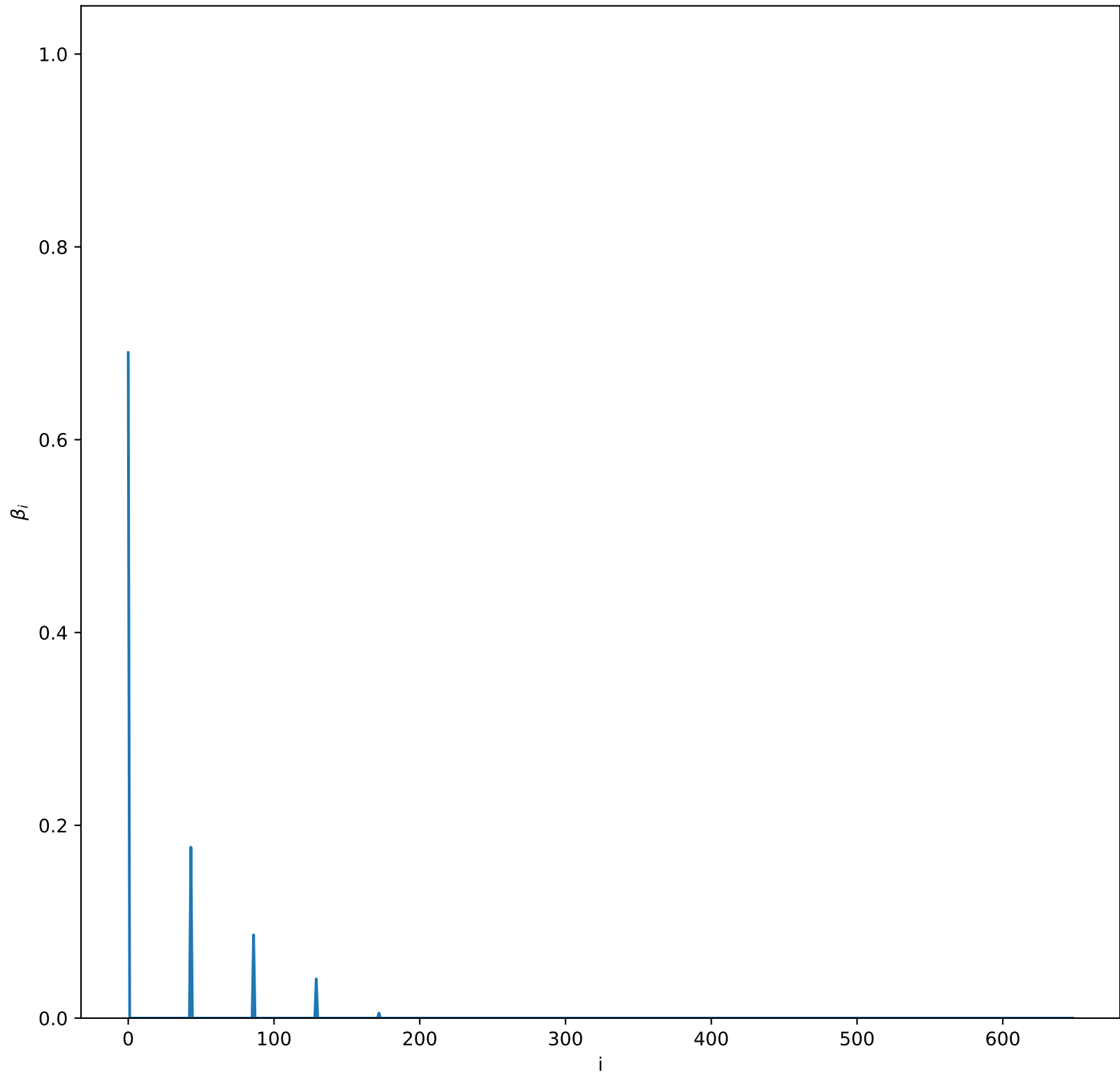
$\mu = 1.55$



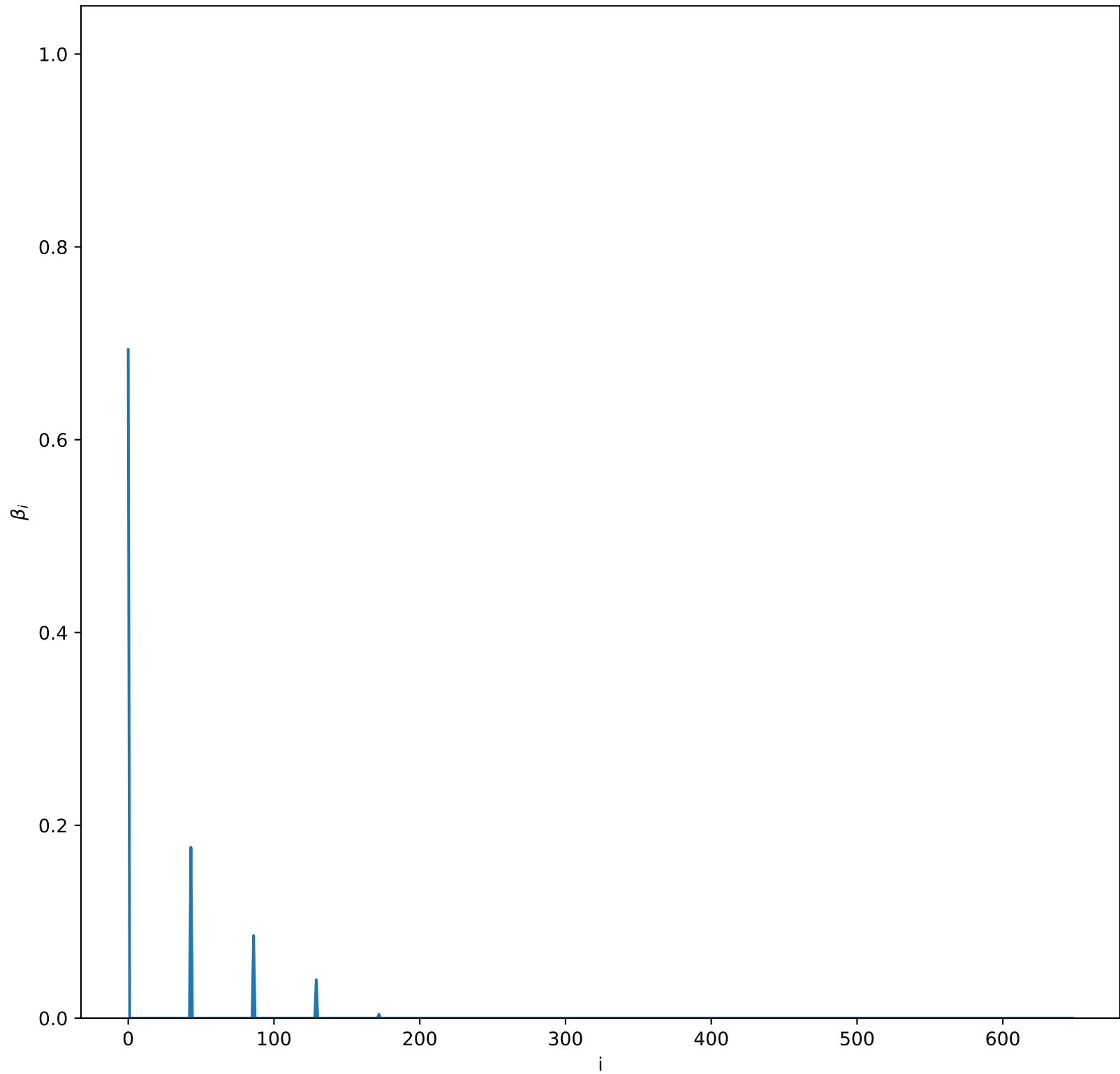
$\mu = 1.56$



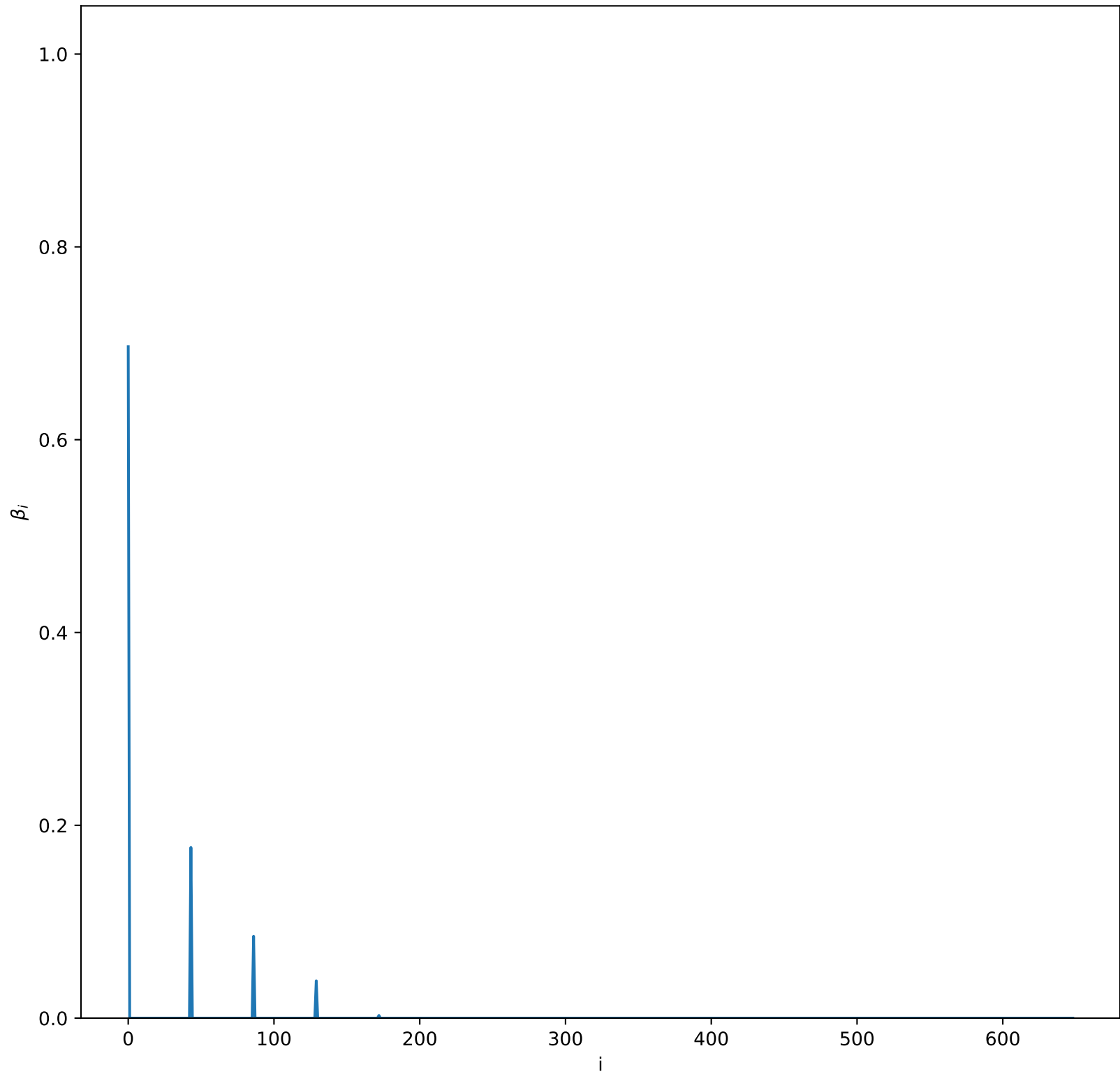
$\mu = 1.57$



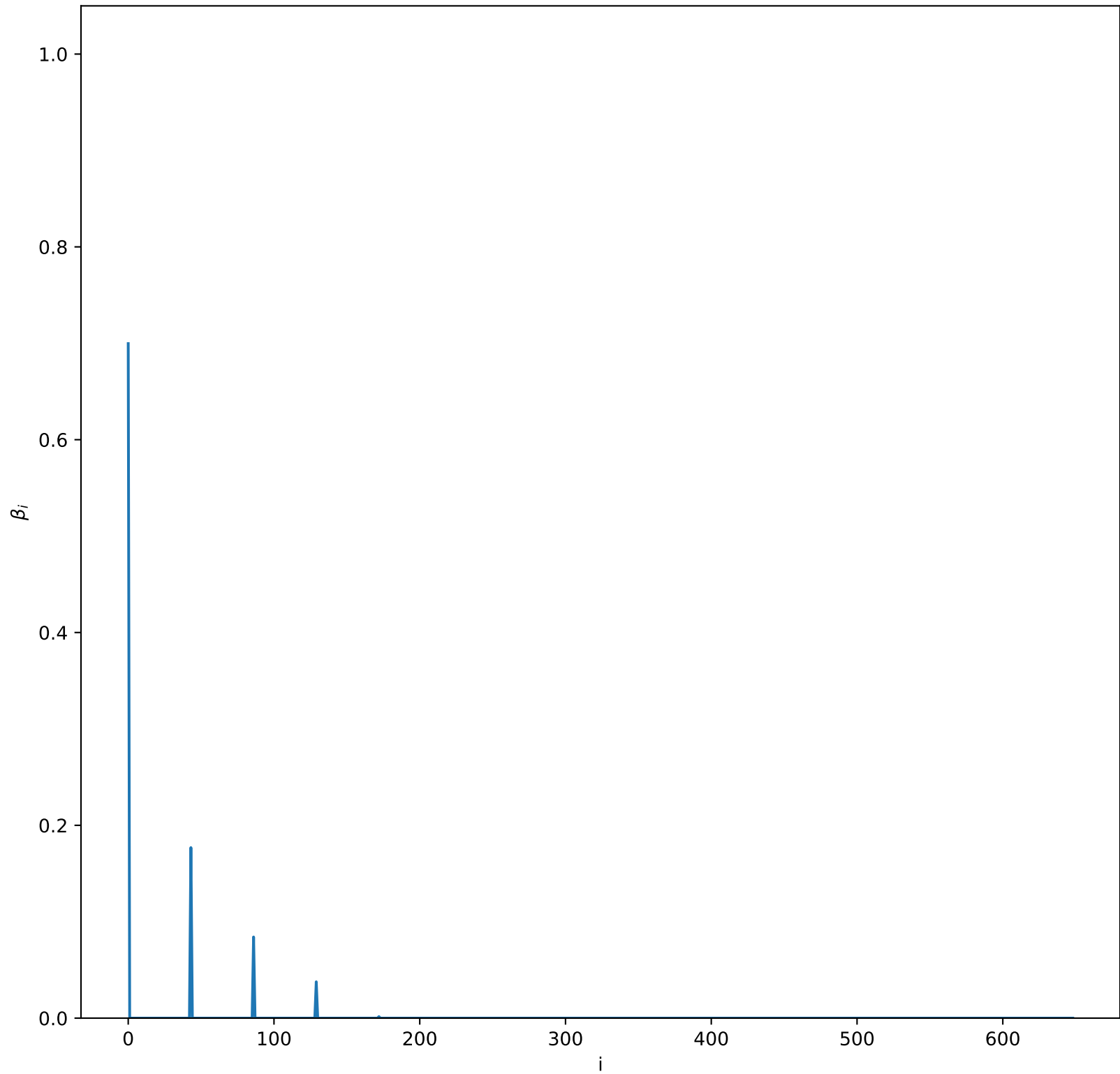
$\mu = 1.58$



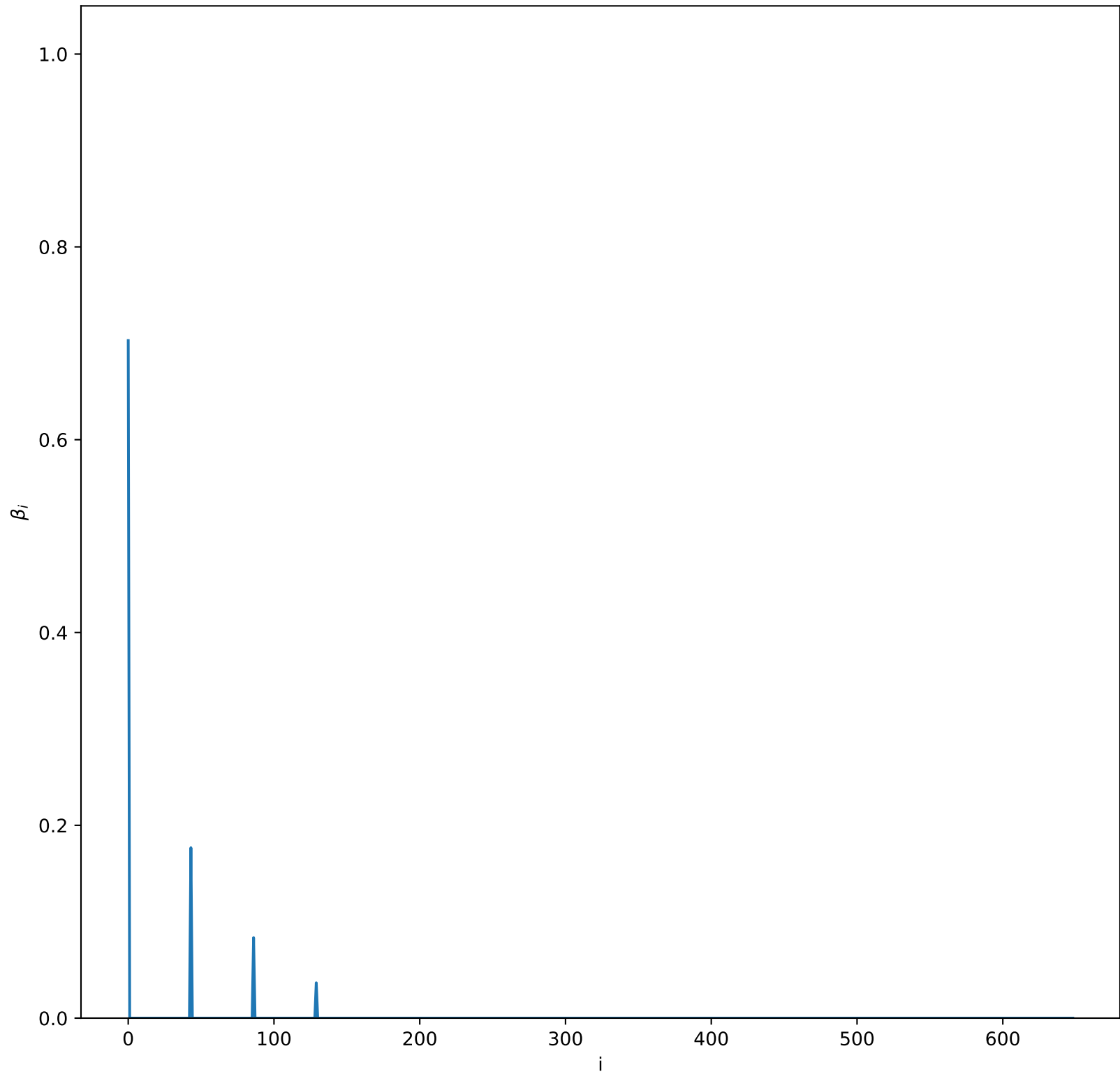
$\mu = 1.59$



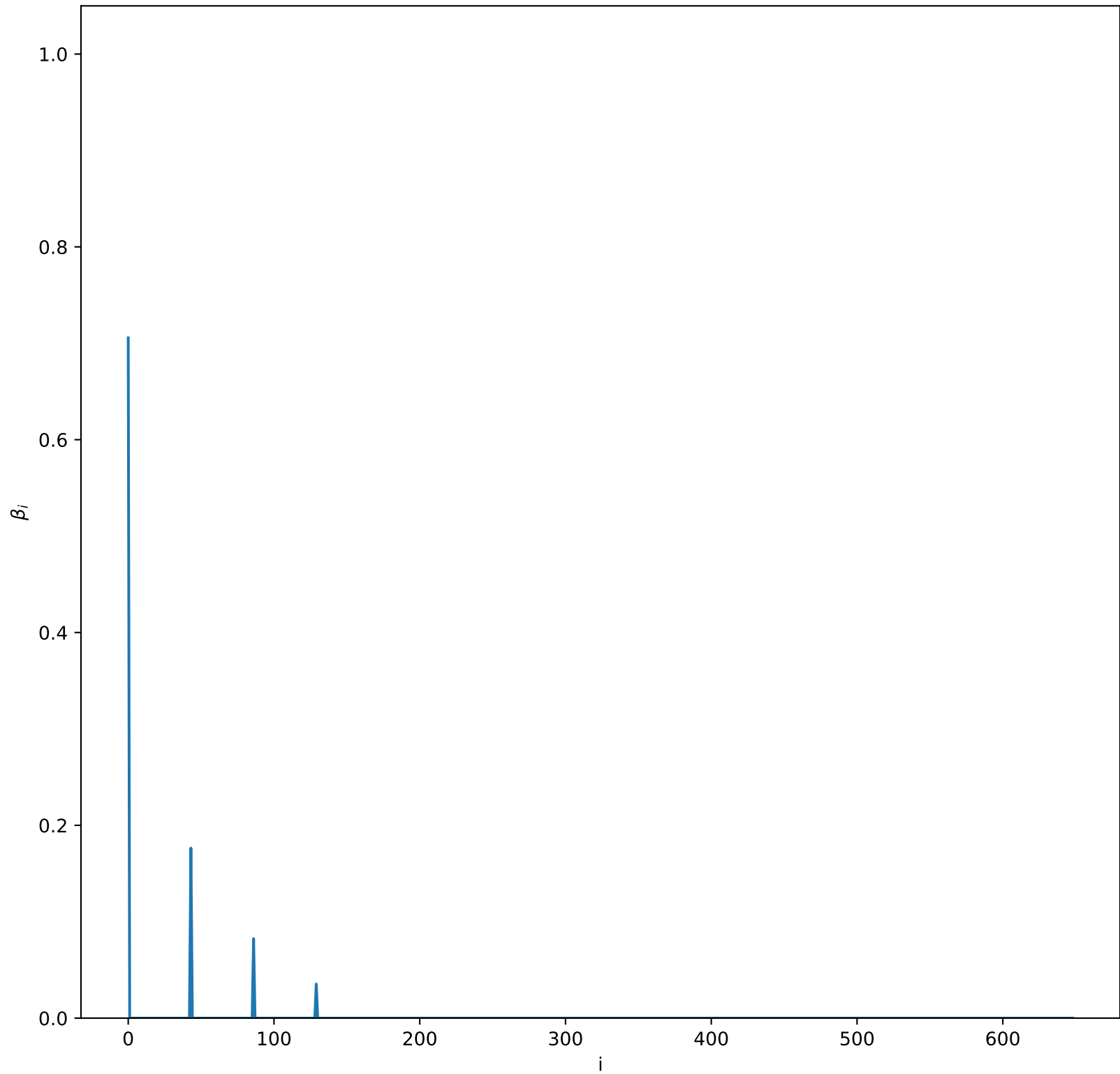
$\mu = 1.60$



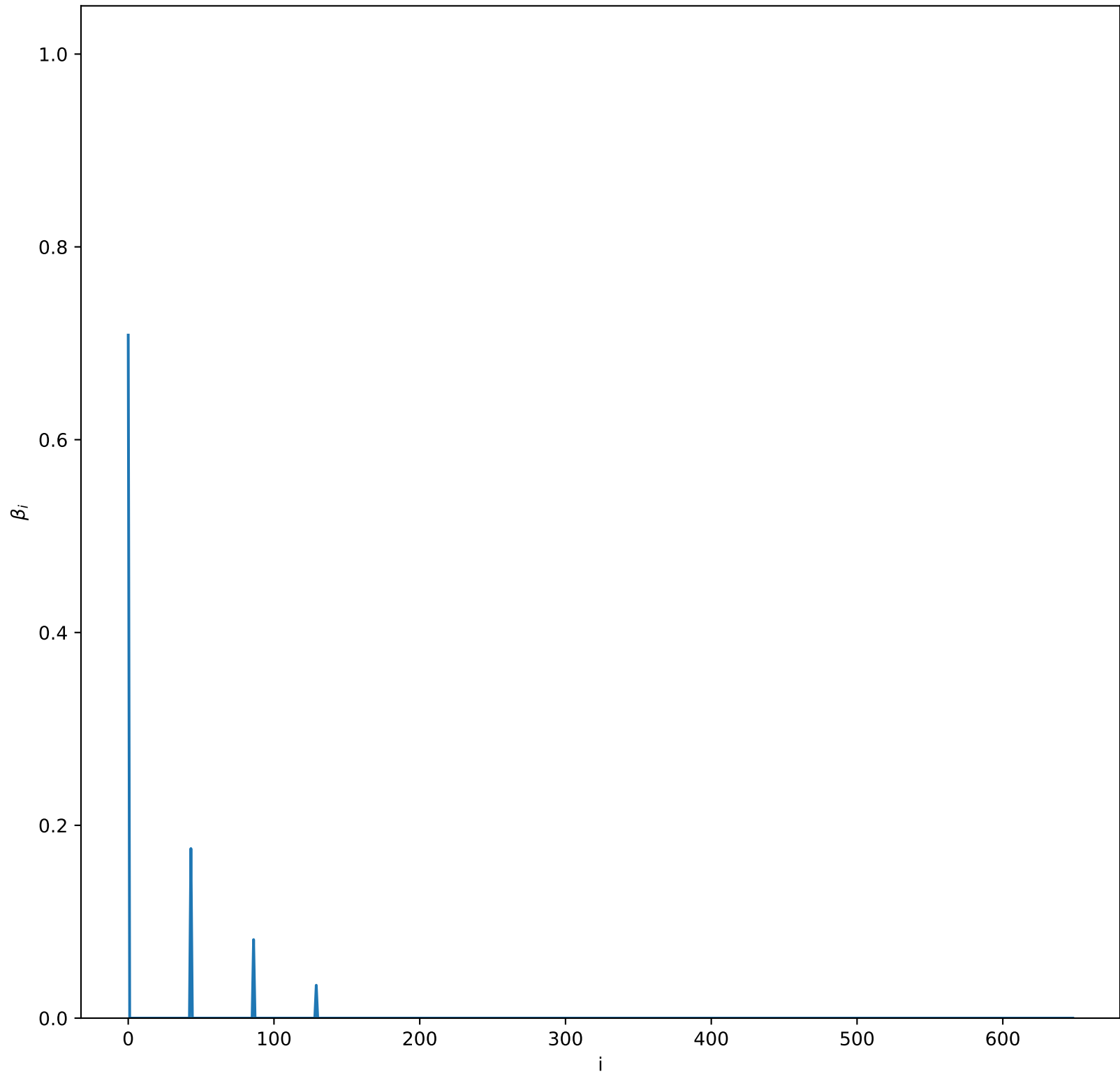
$\mu = 1.61$



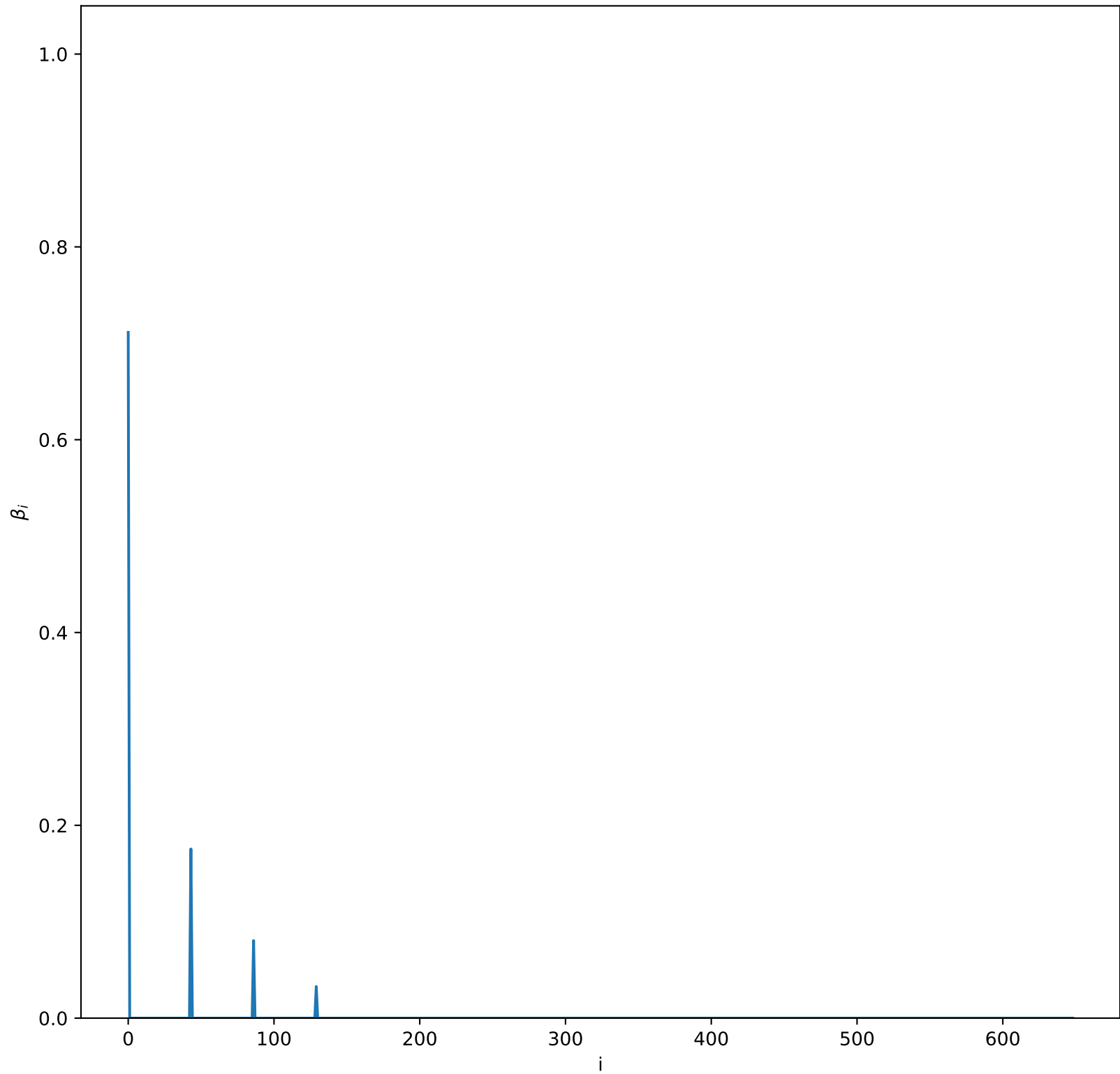
$\mu = 1.62$



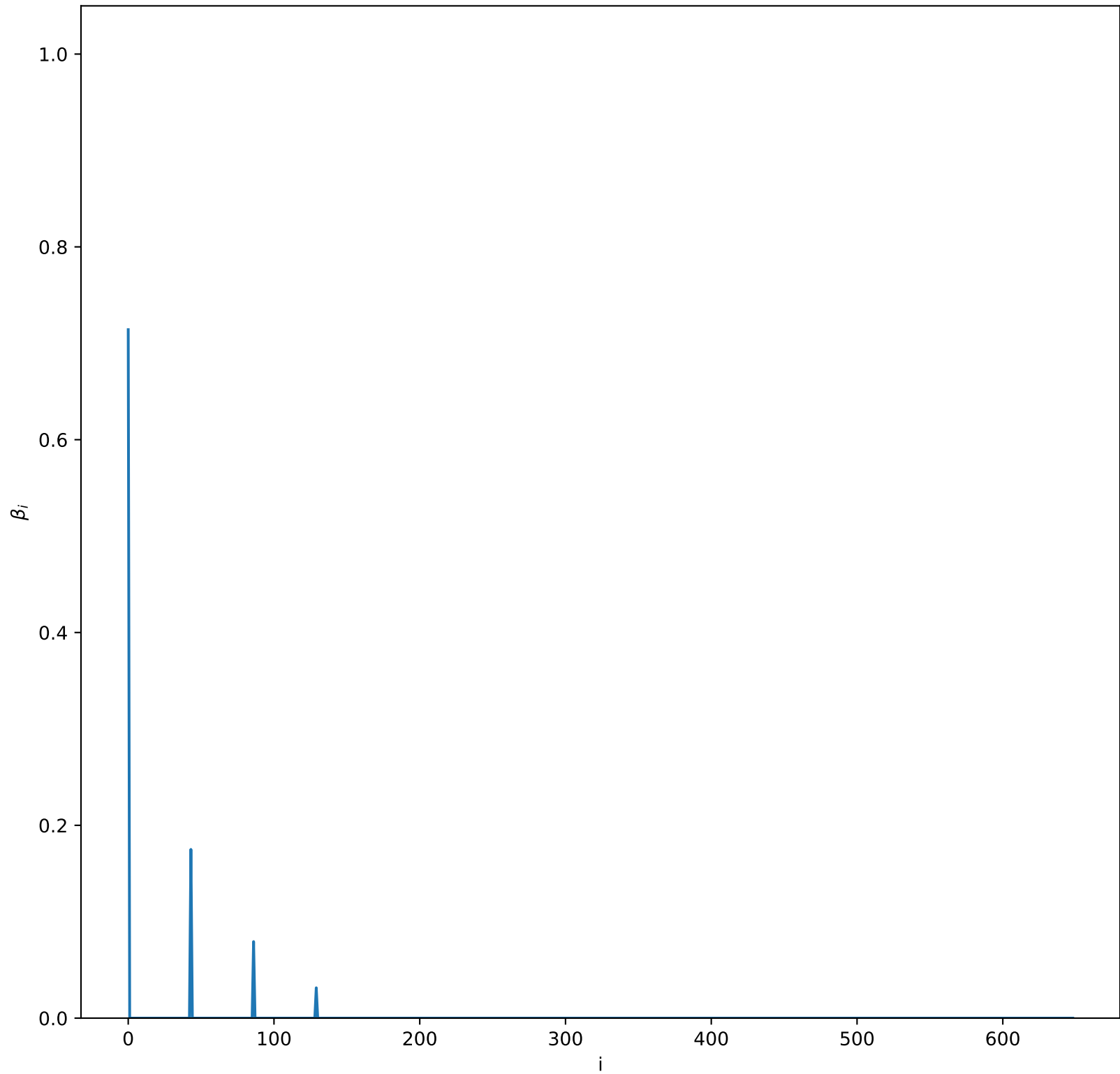
$\mu = 1.63$



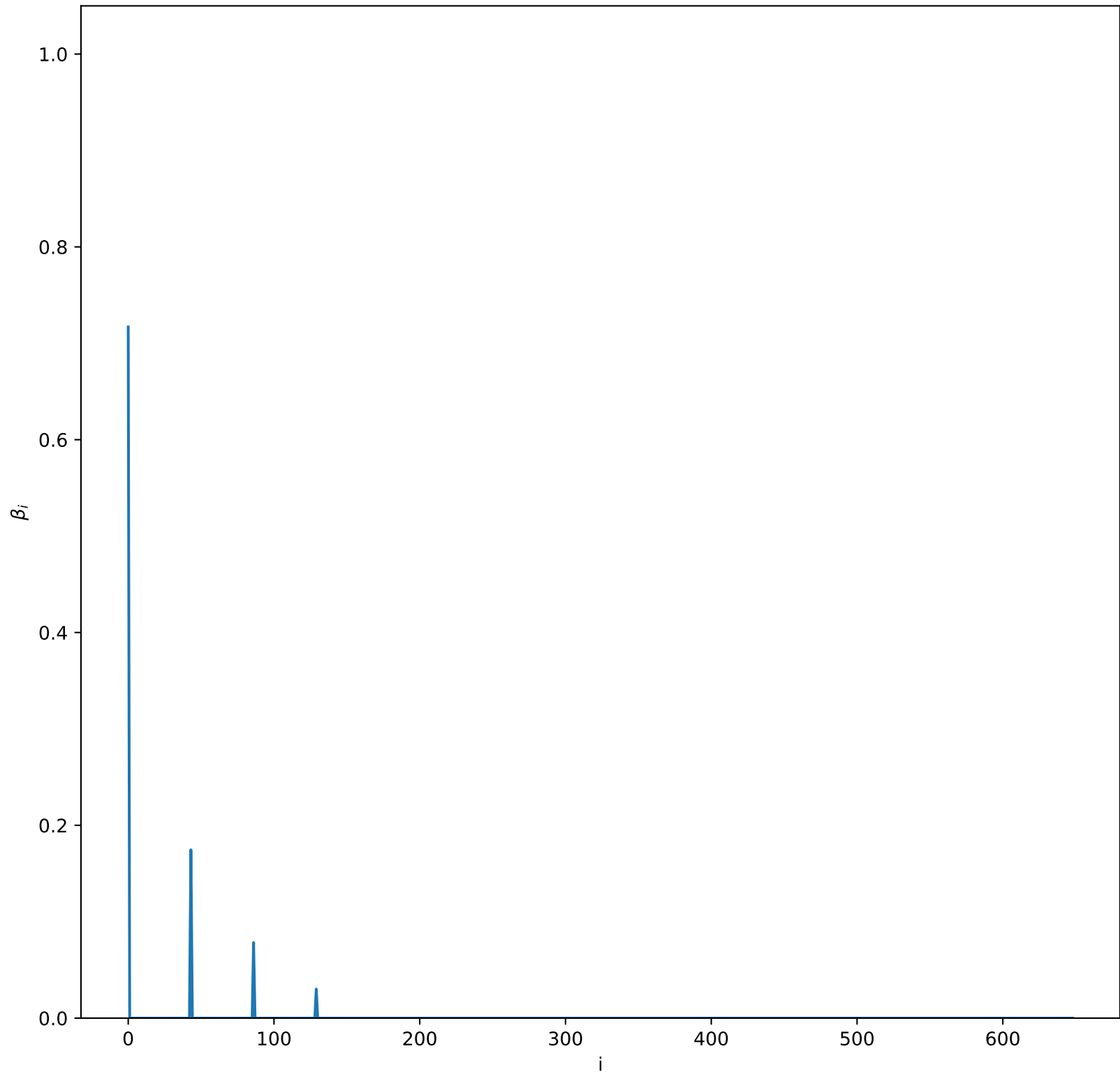
$\mu = 1.64$



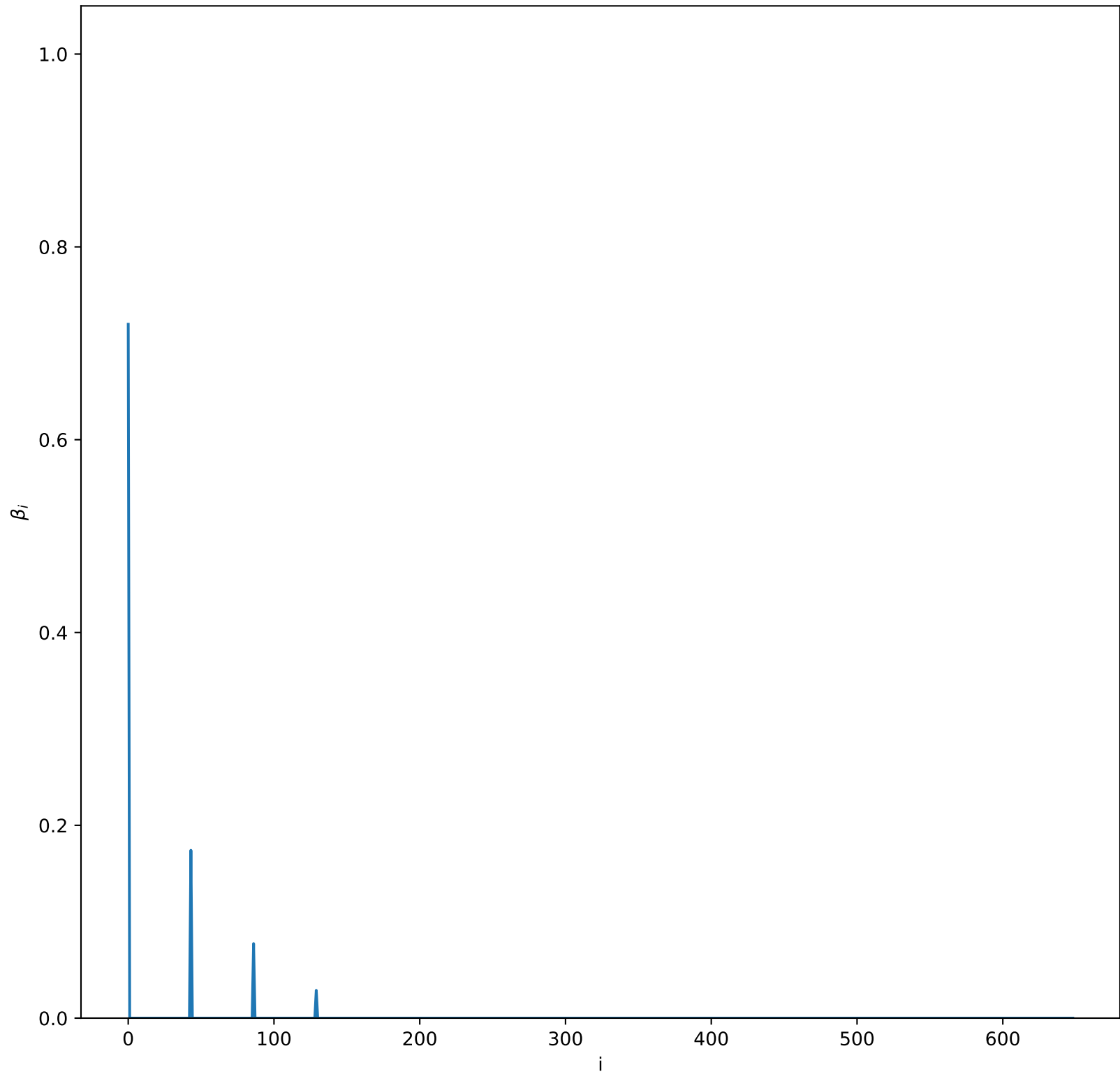
$\mu = 1.65$



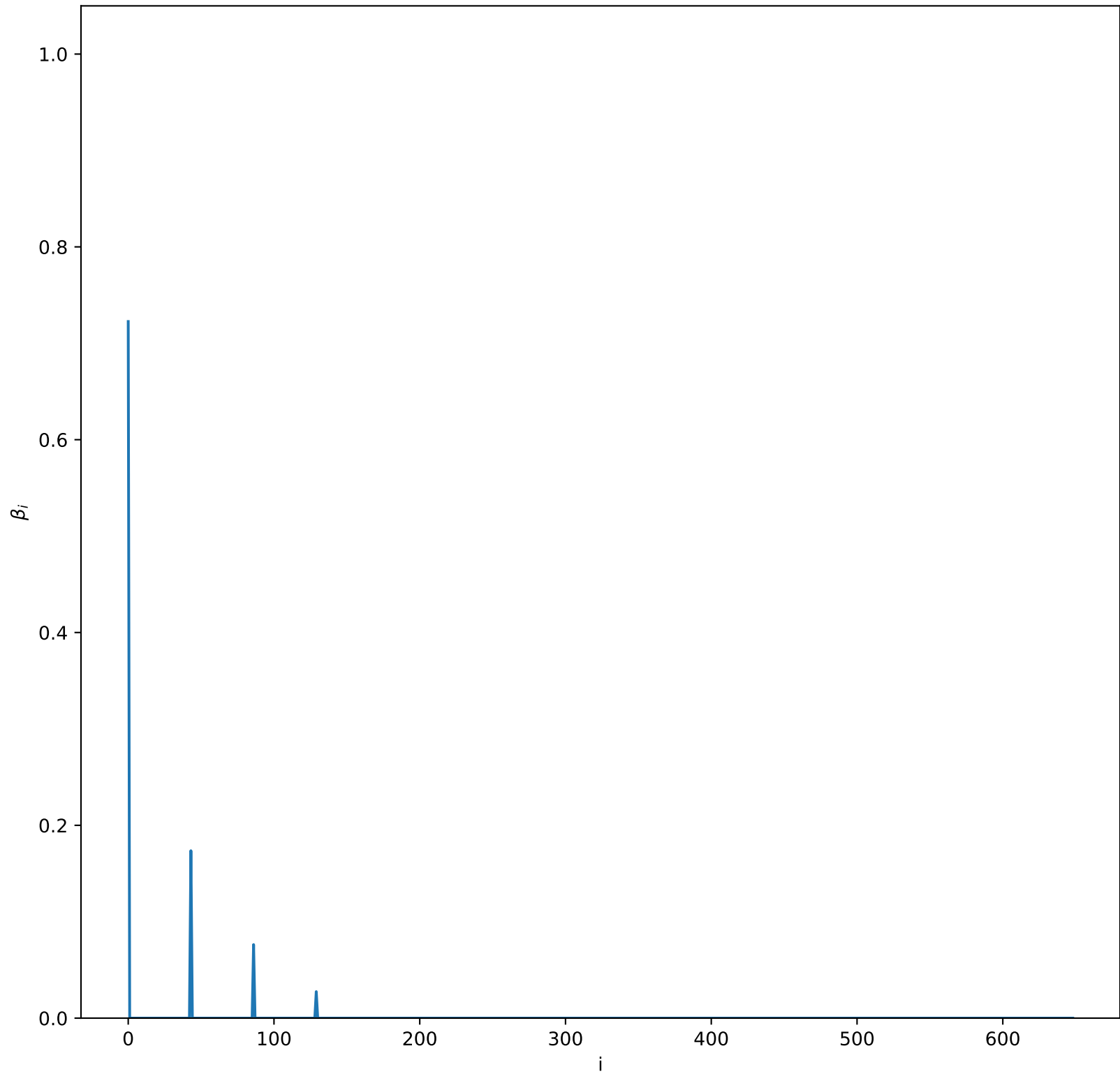
$\mu = 1.66$



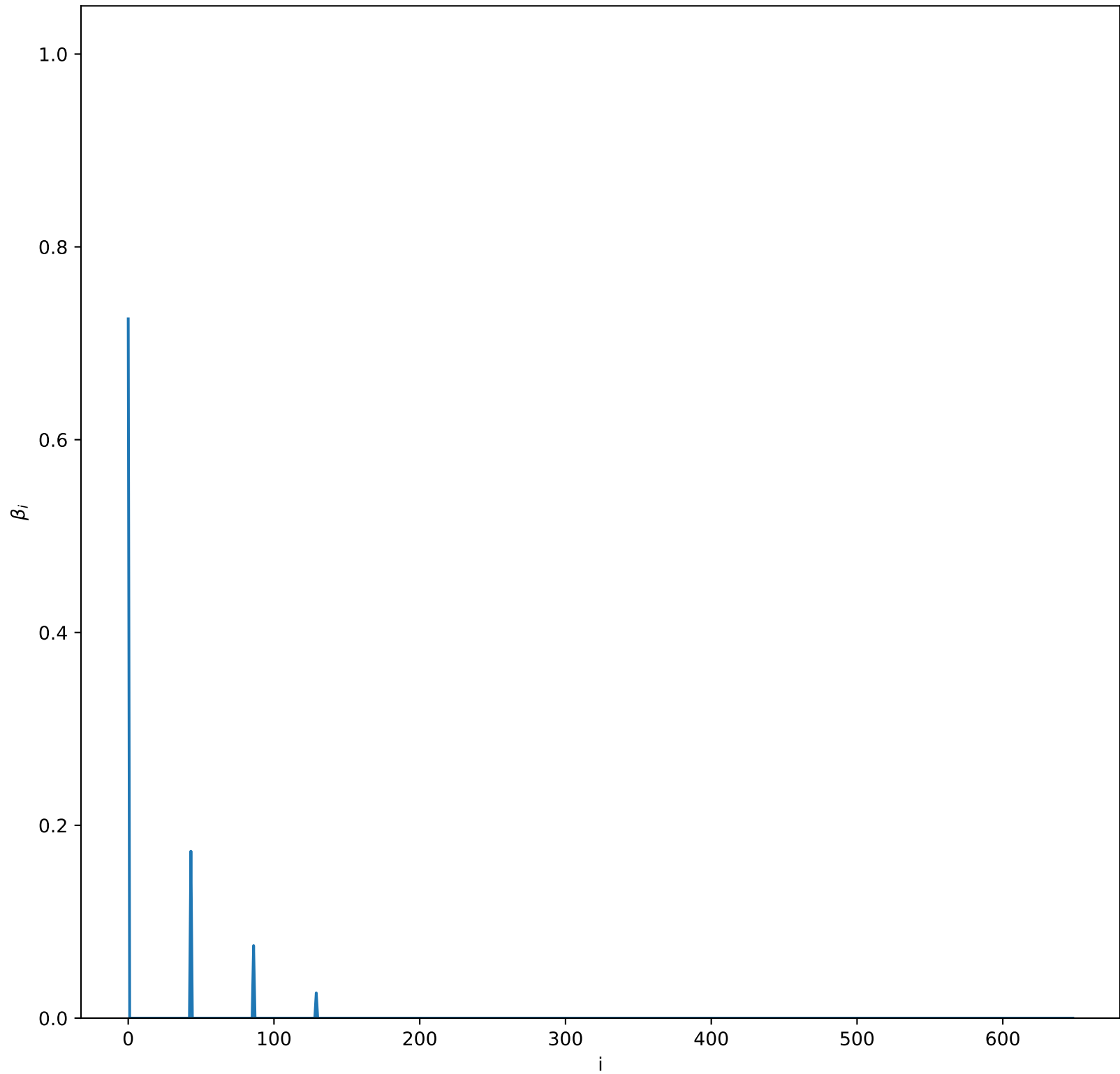
$\mu = 1.67$



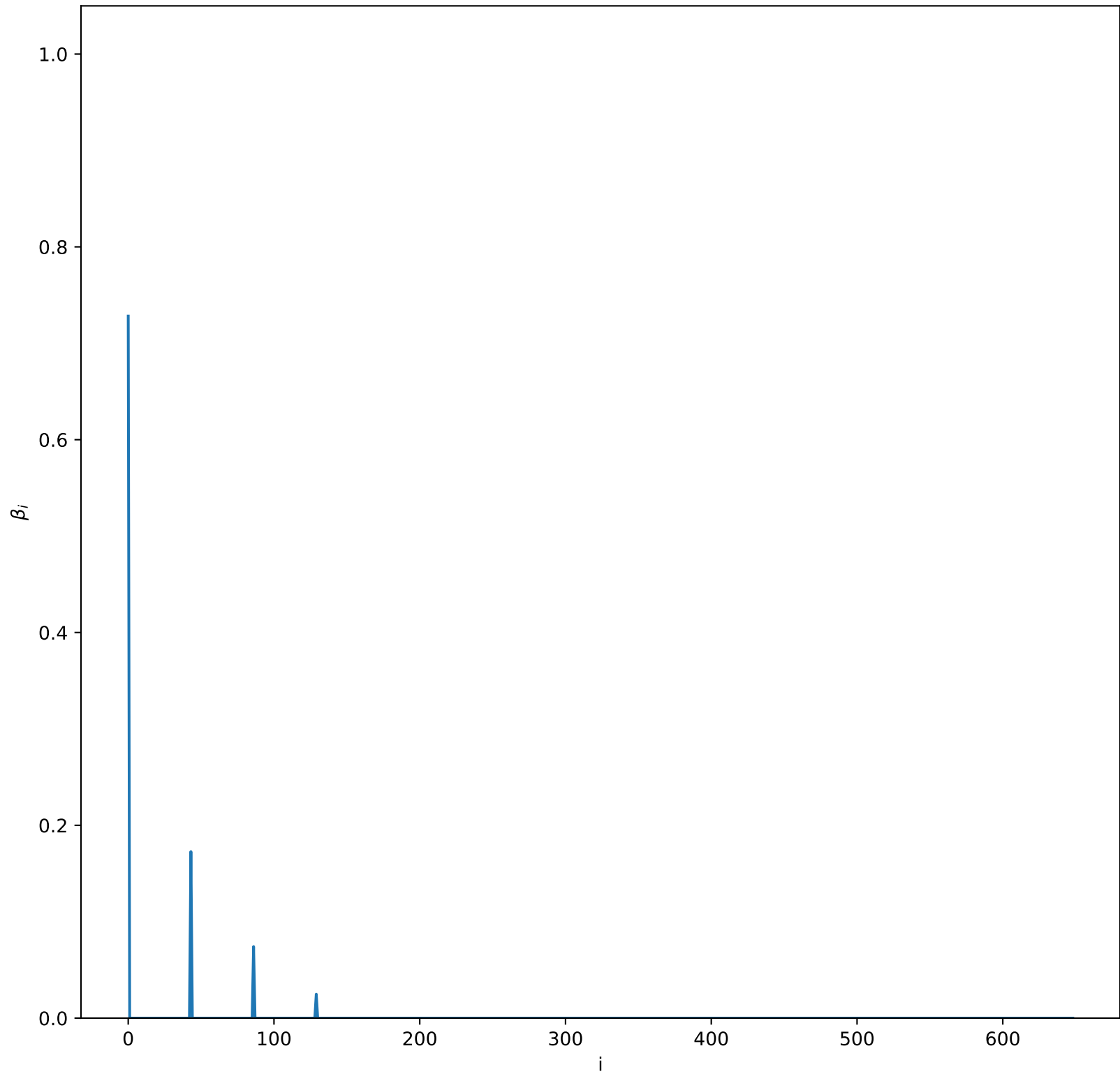
$\mu = 1.68$



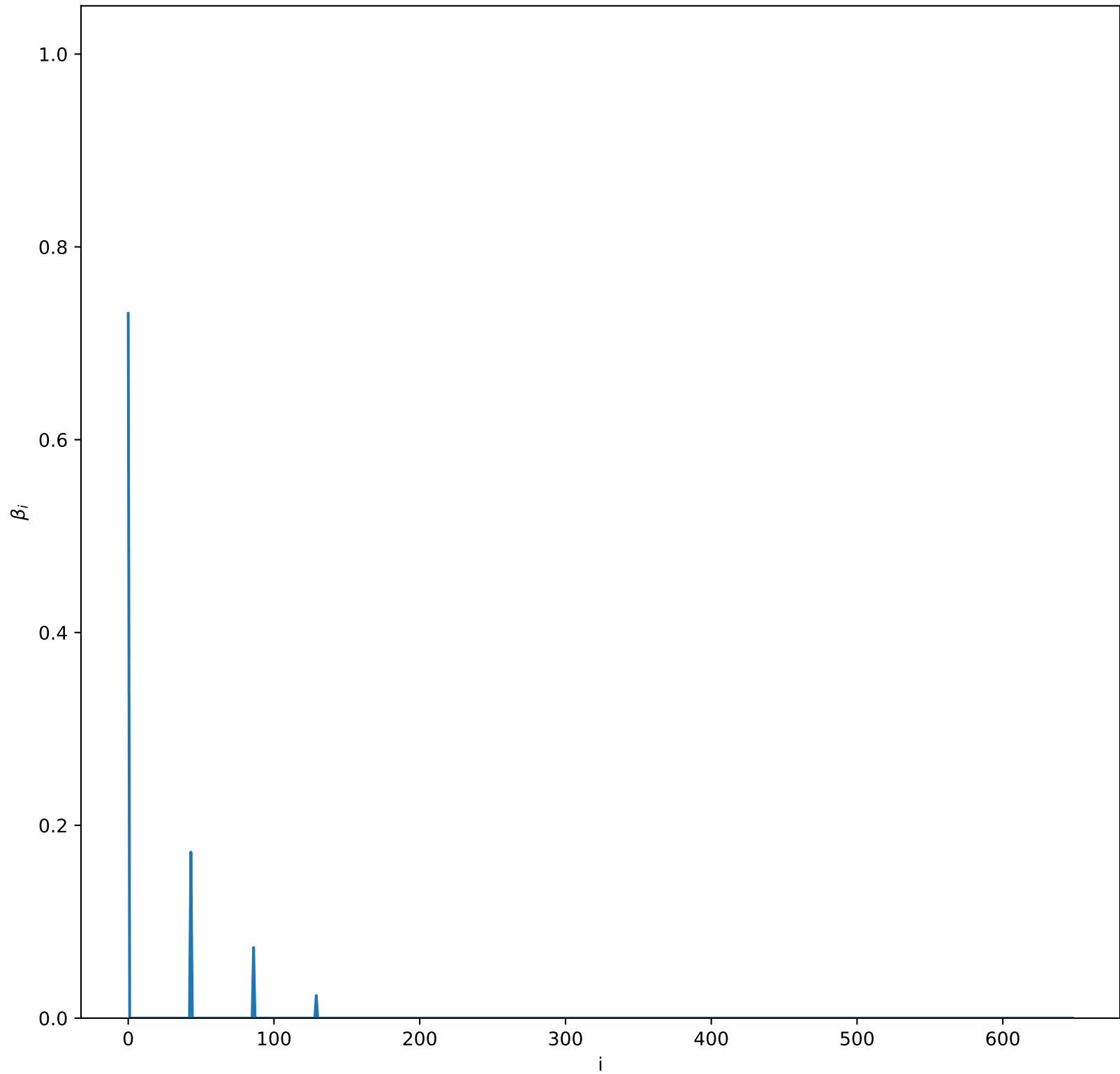
$\mu = 1.69$



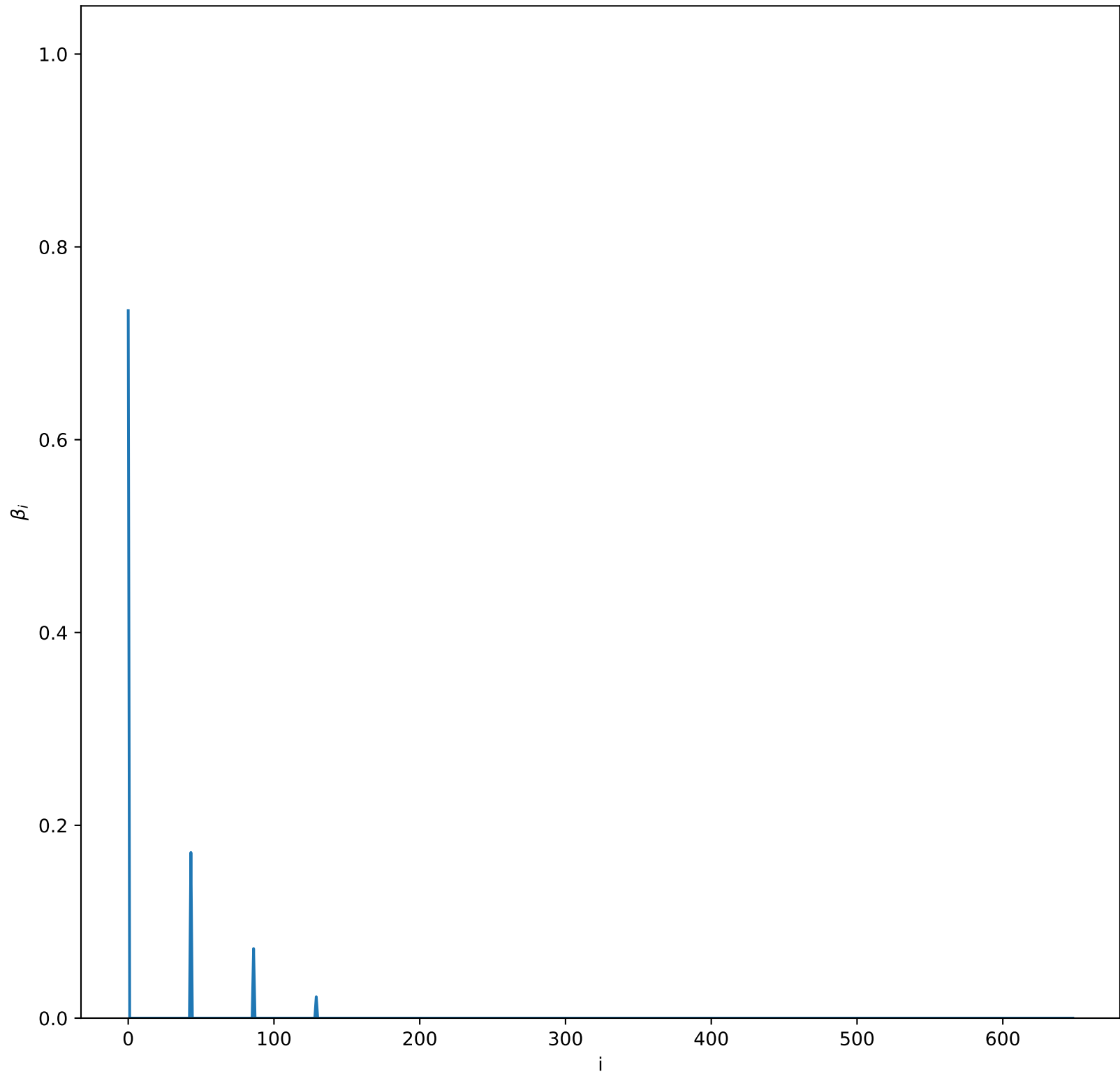
$\mu = 1.70$



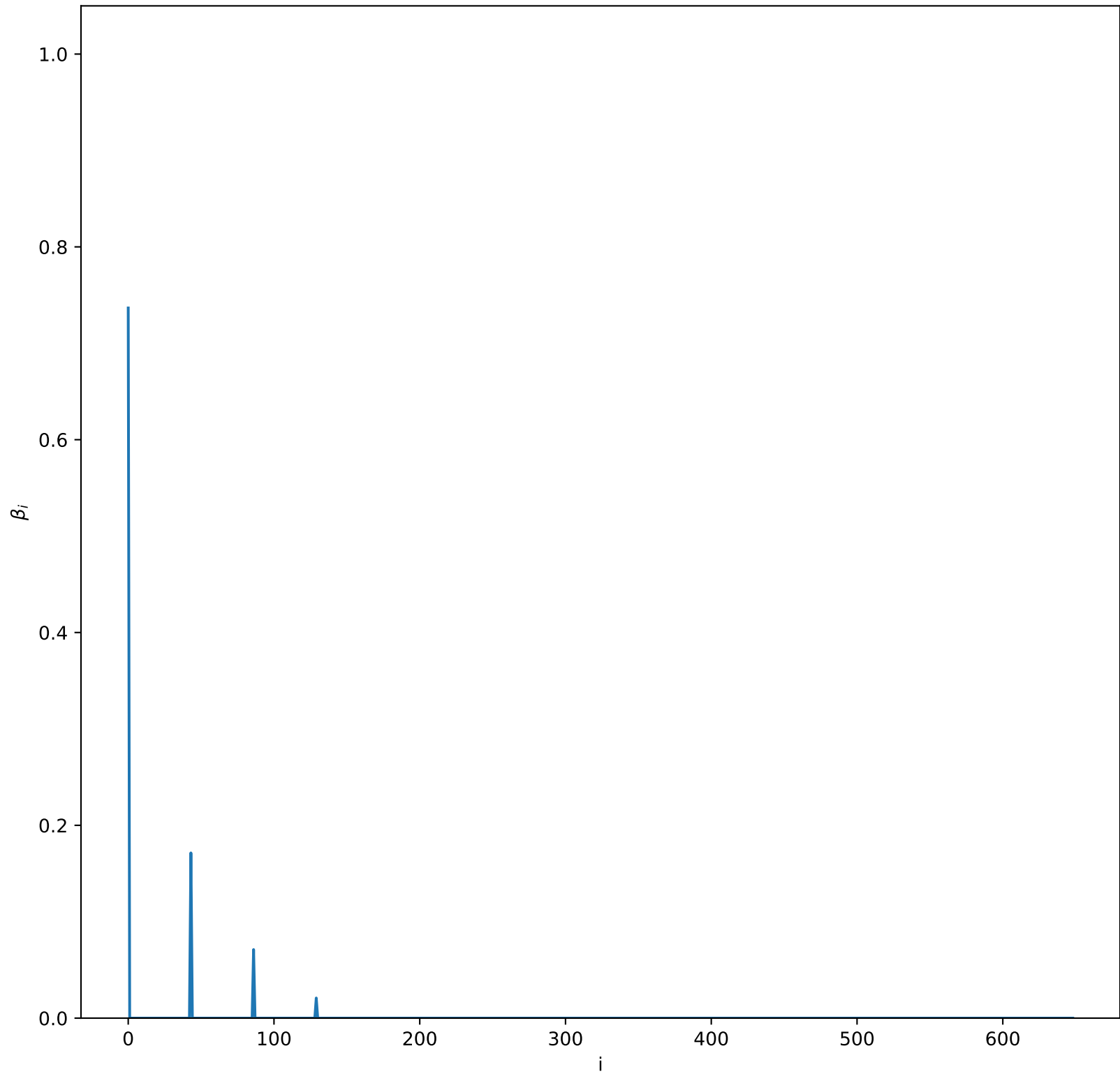
$\mu = 1.71$



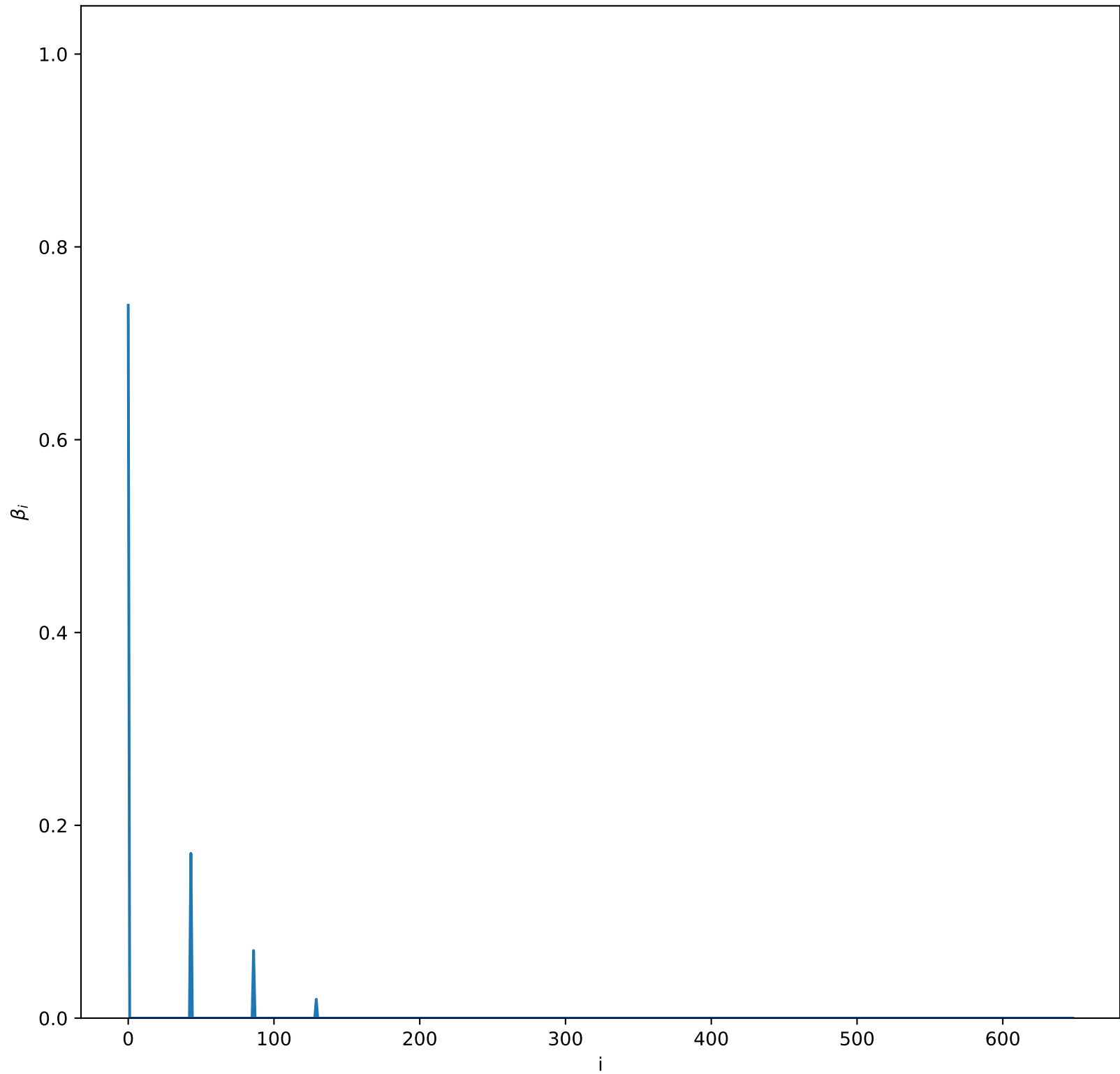
$\mu = 1.72$



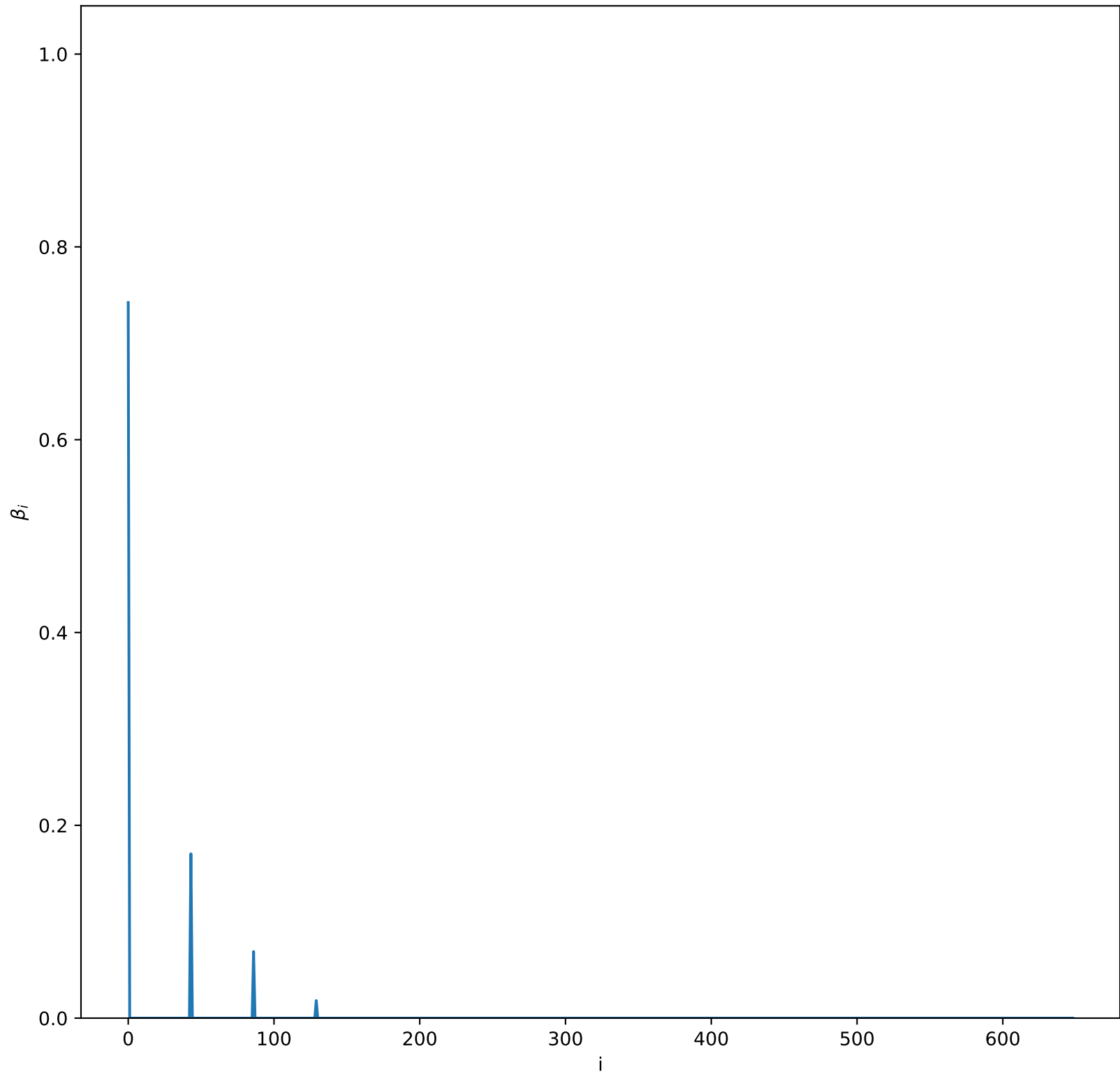
$\mu = 1.73$



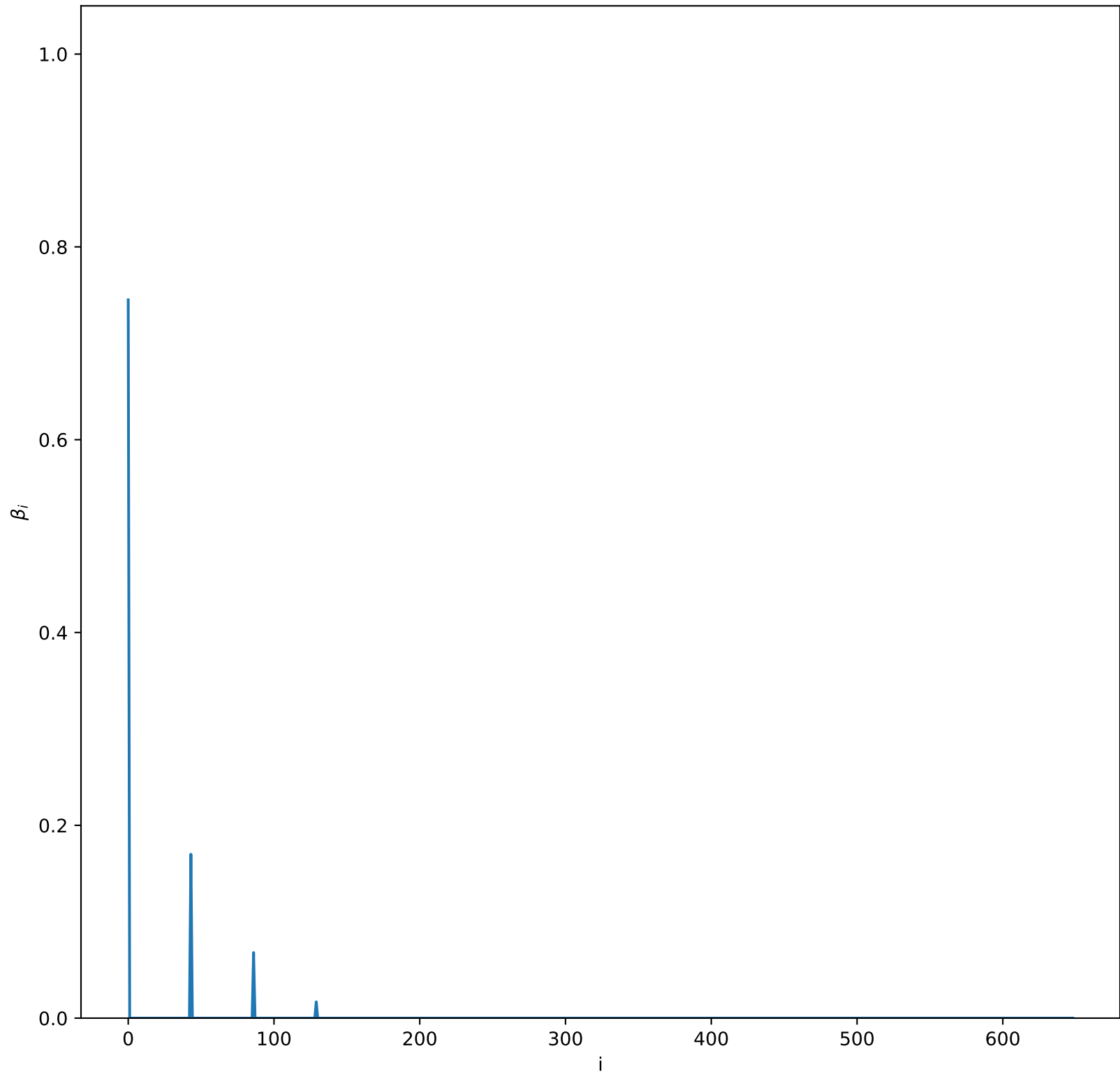
$\mu = 1.74$



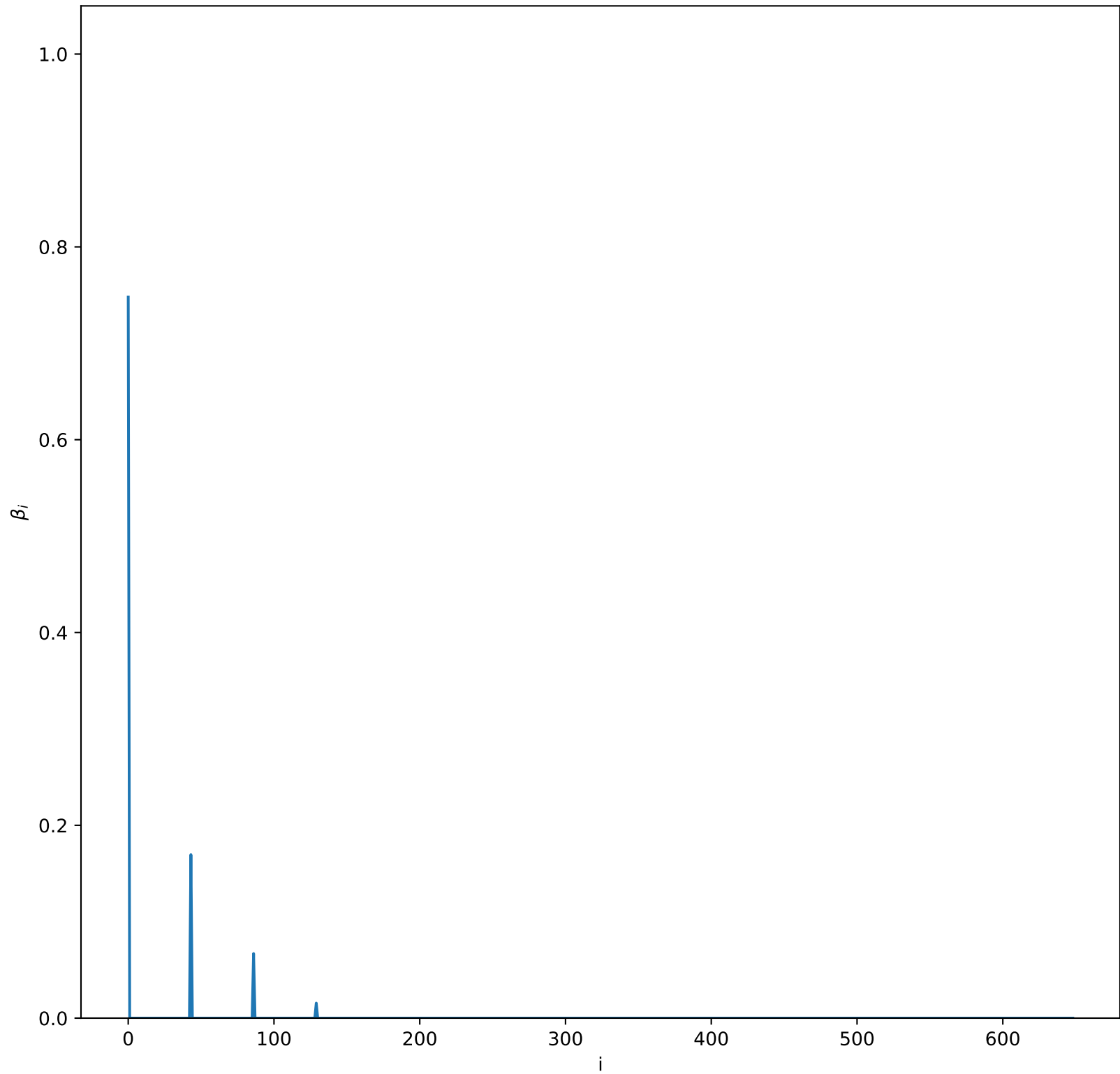
$\mu = 1.75$



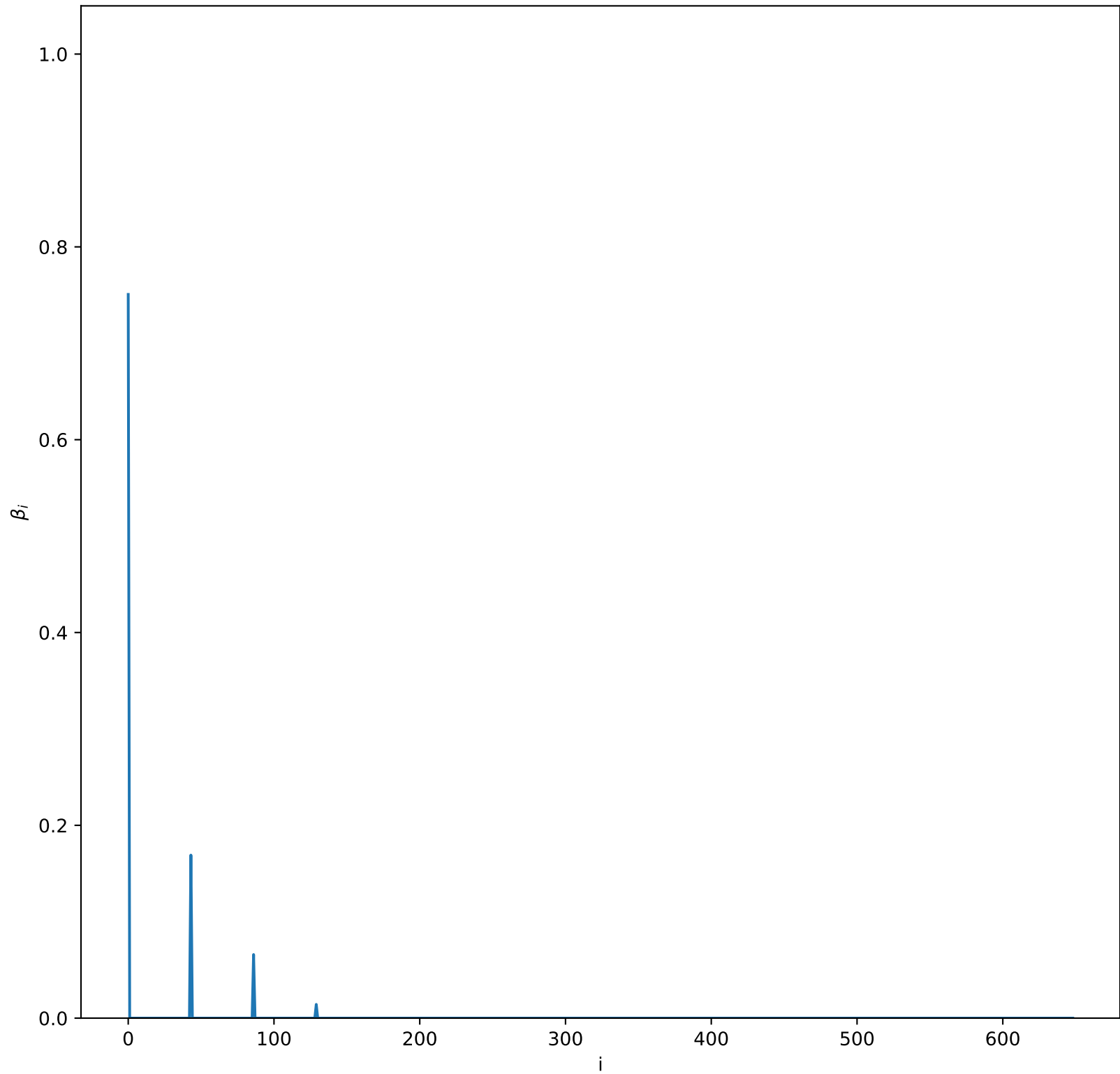
$\mu = 1.76$



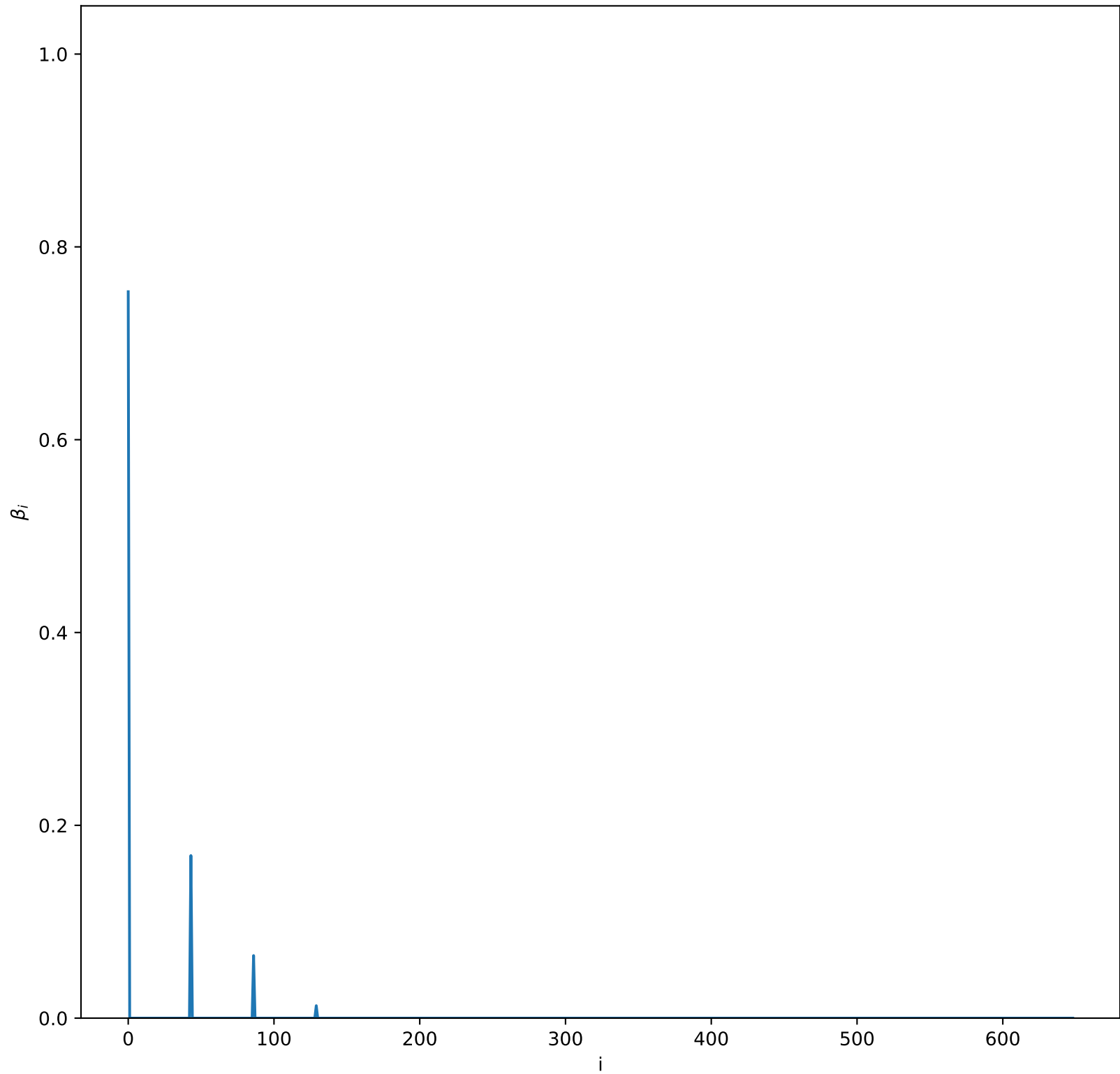
$\mu = 1.77$



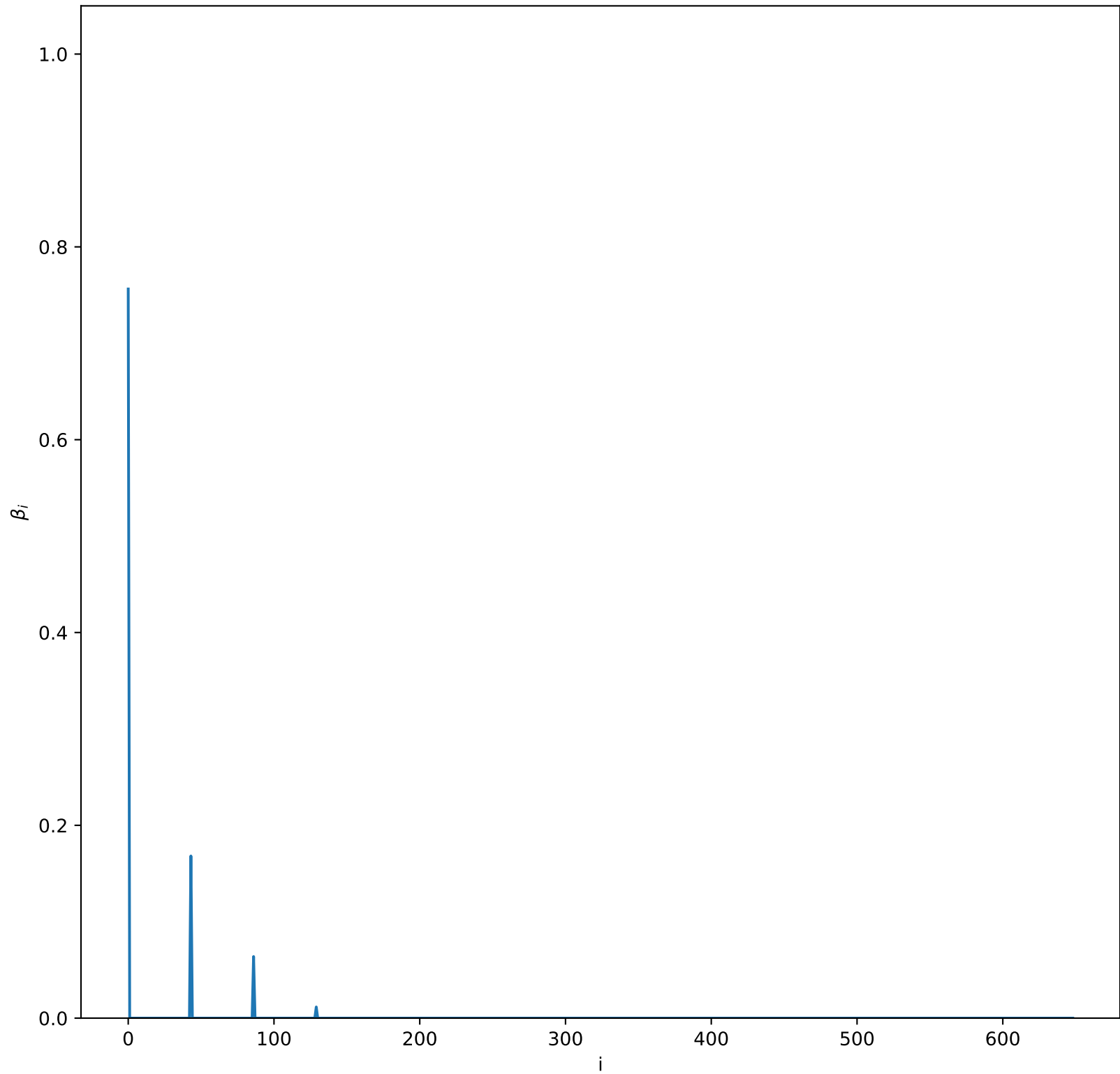
$\mu = 1.78$



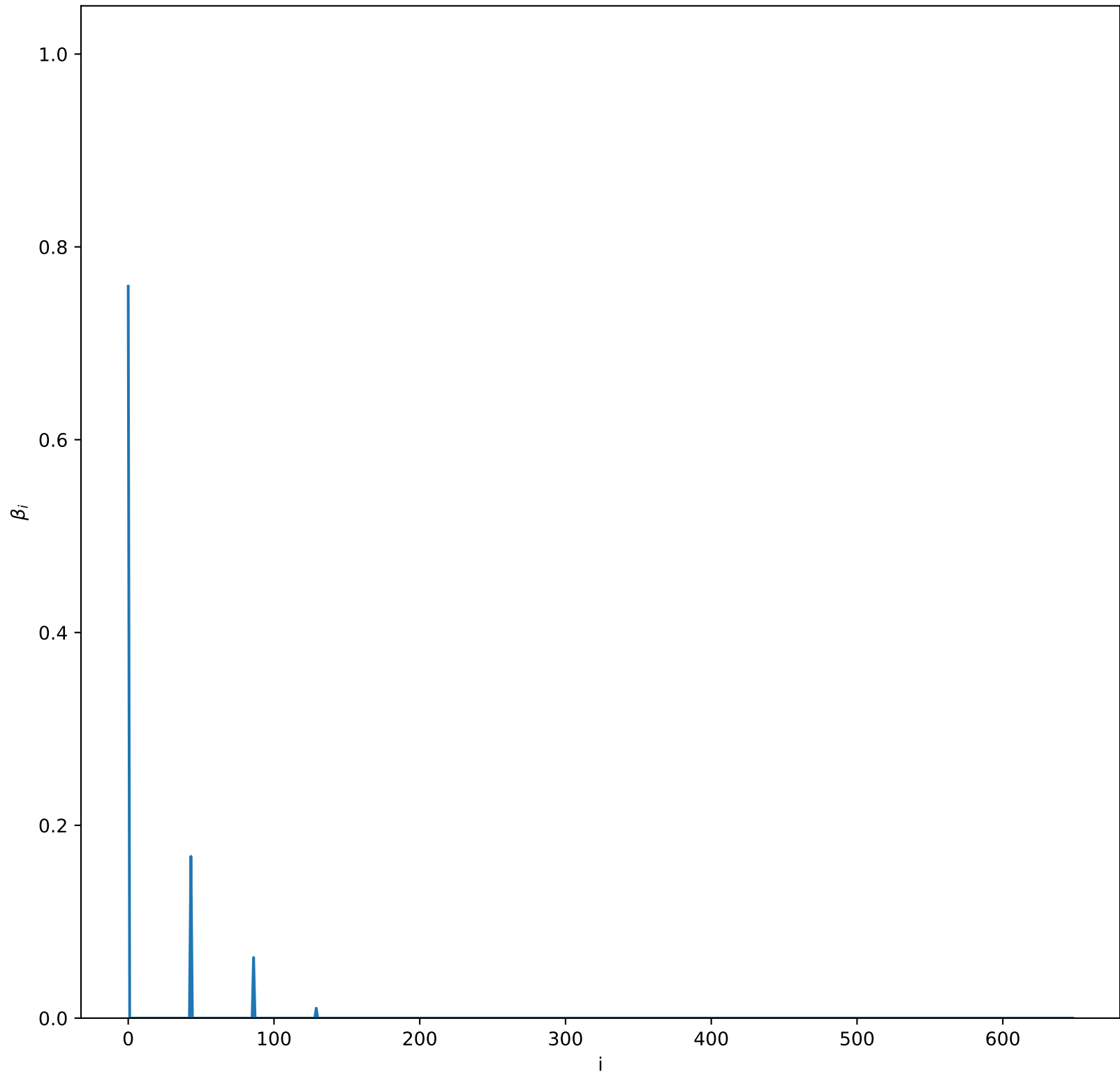
$\mu = 1.79$



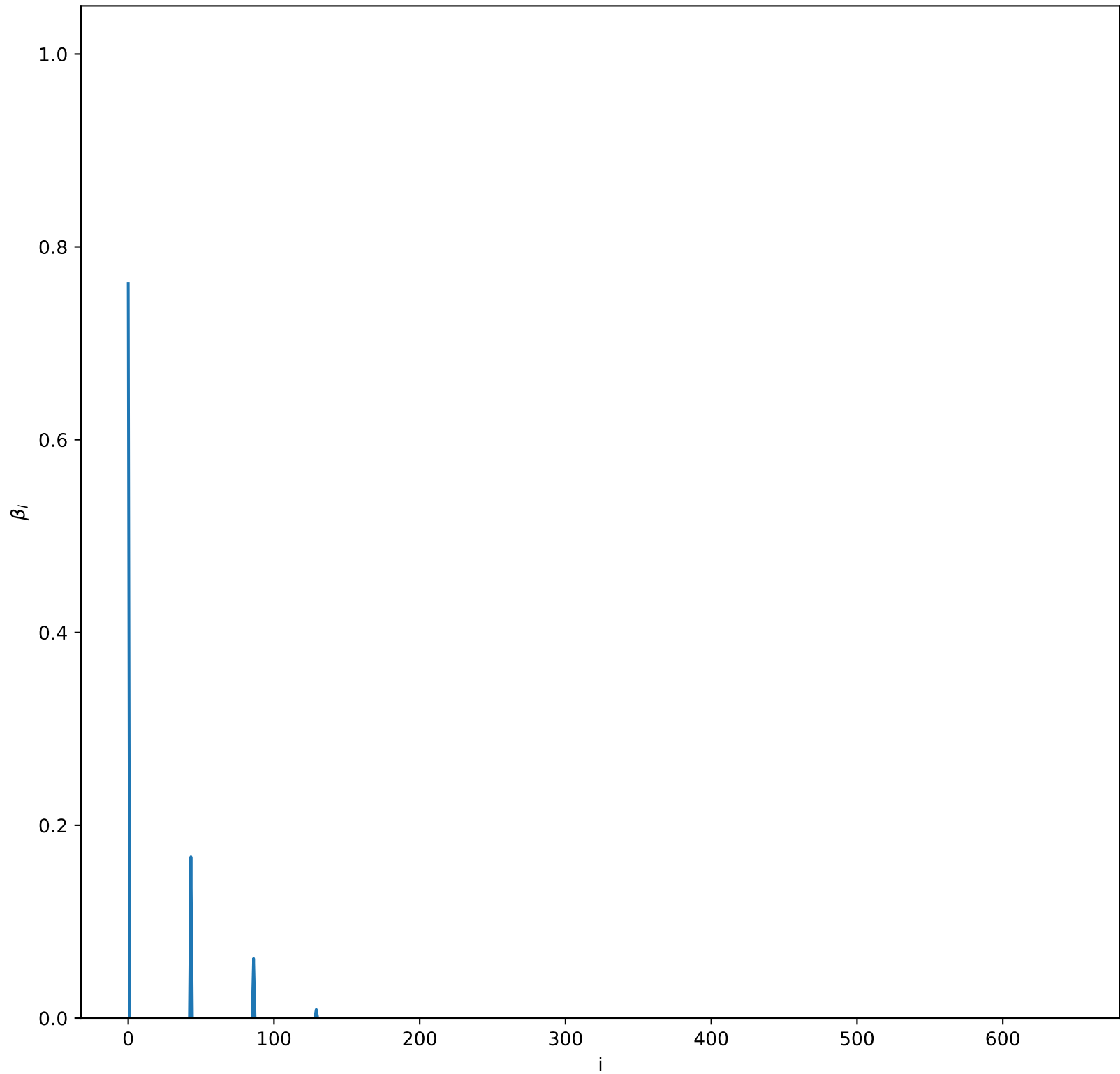
$\mu = 1.80$



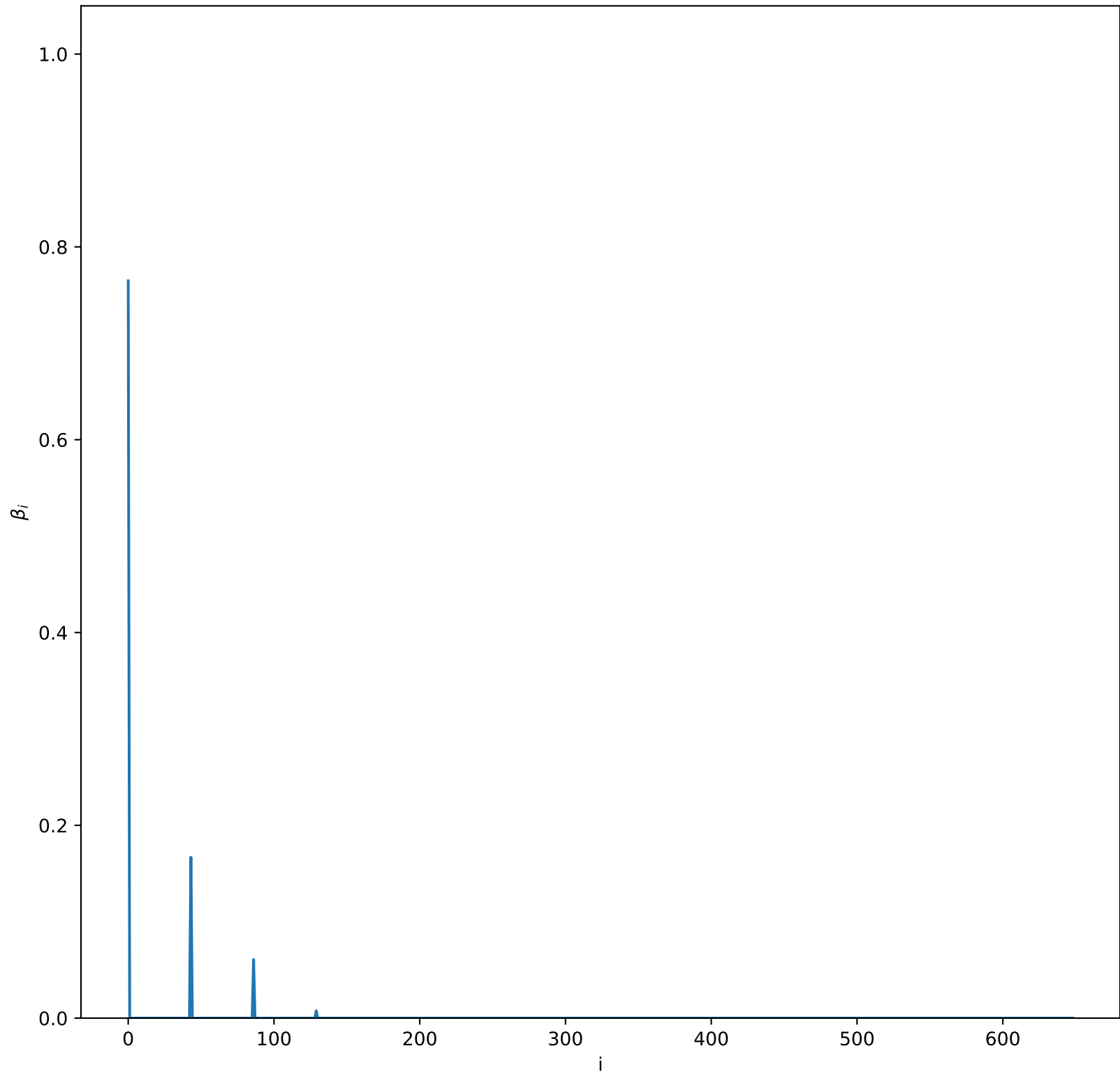
$\mu = 1.81$



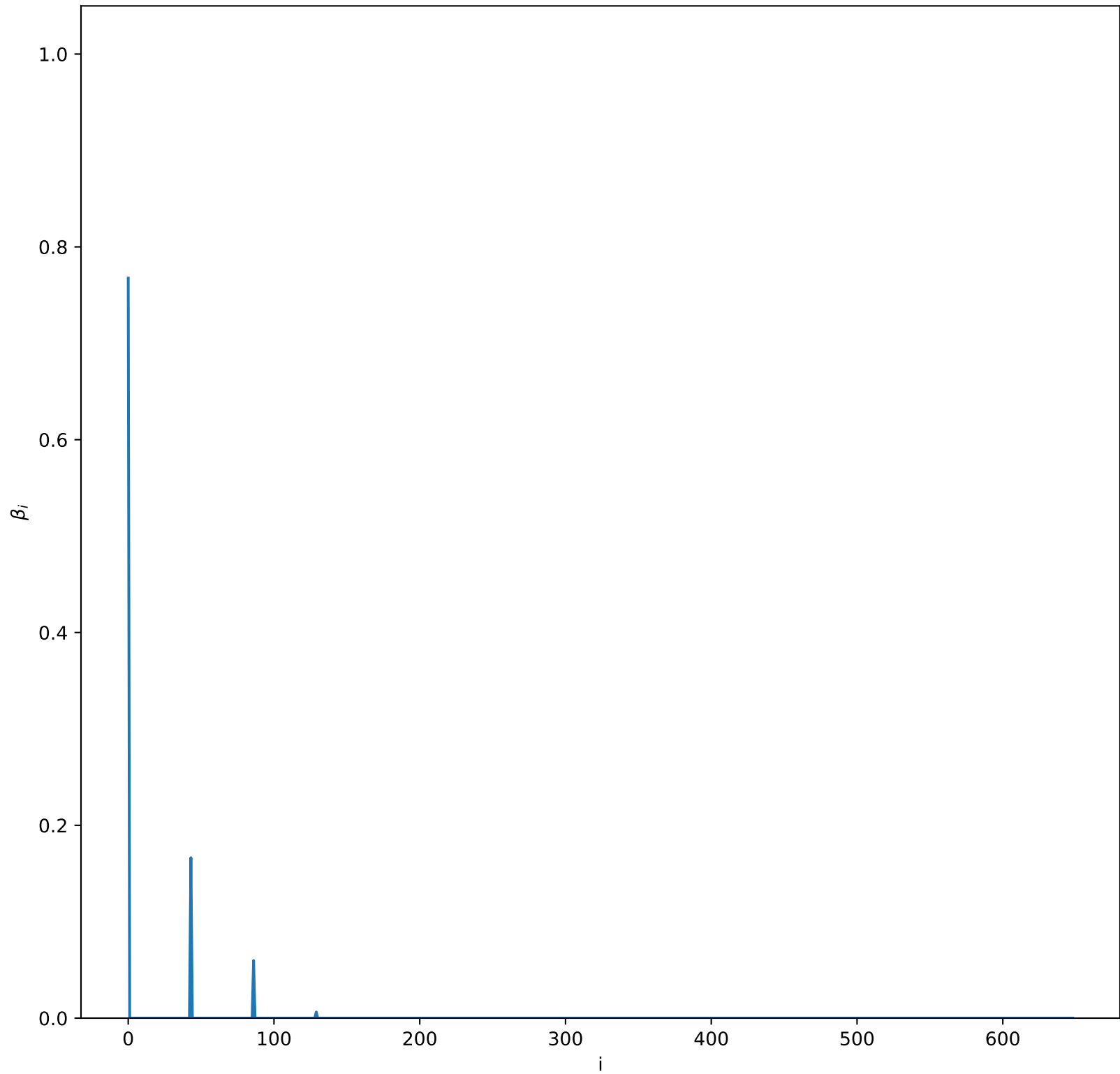
$\mu = 1.82$



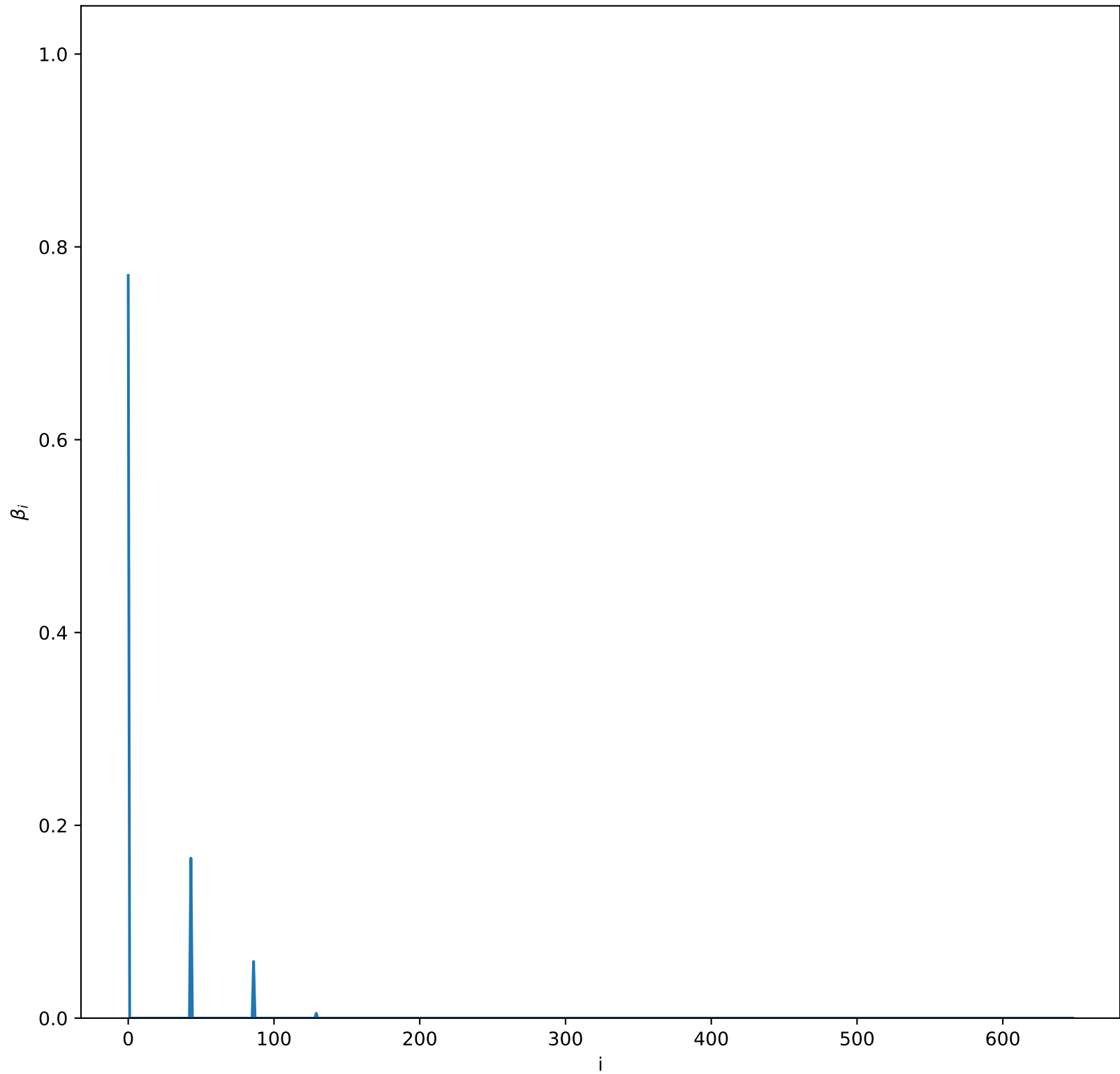
$\mu = 1.83$



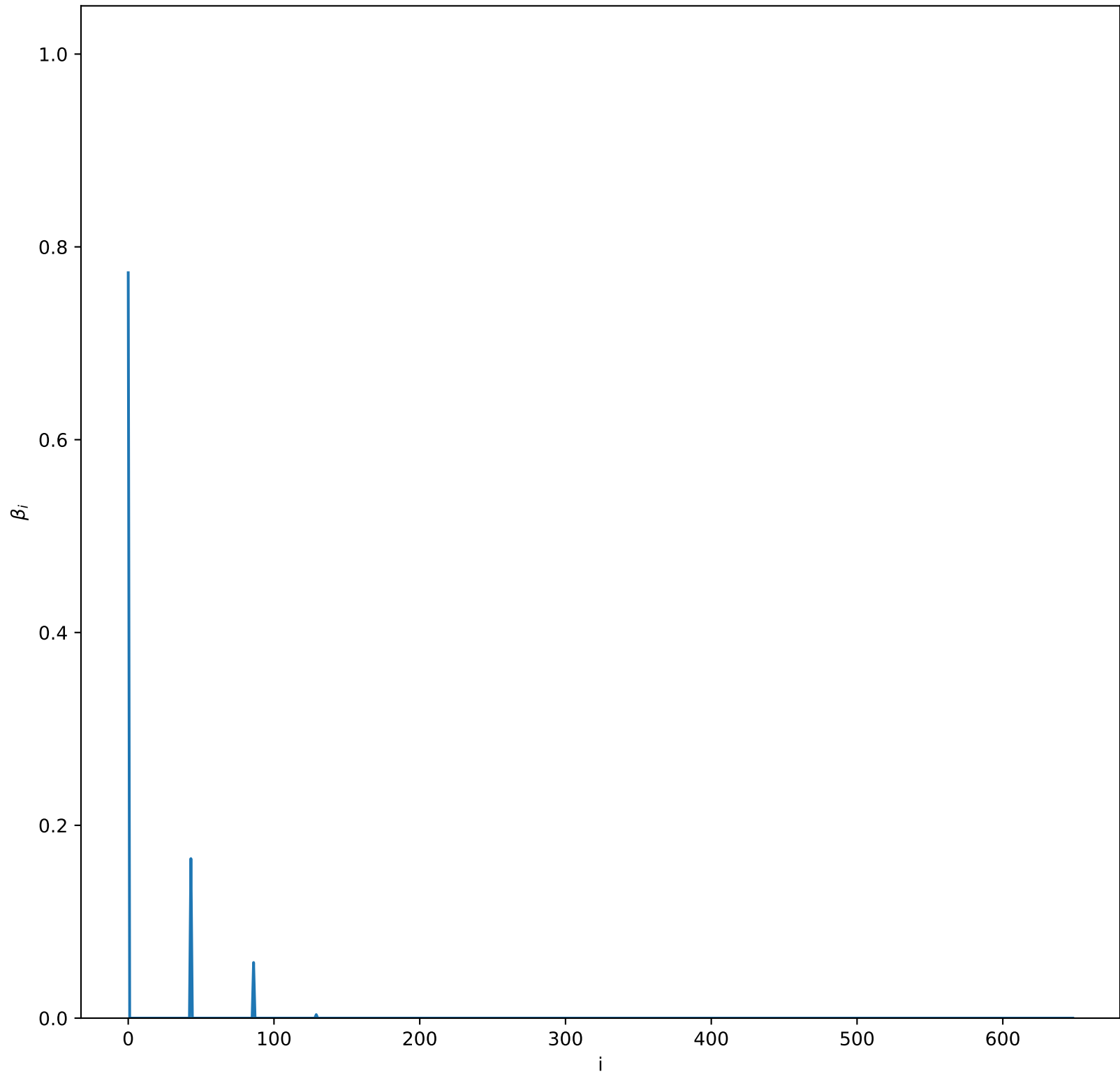
$\mu = 1.84$



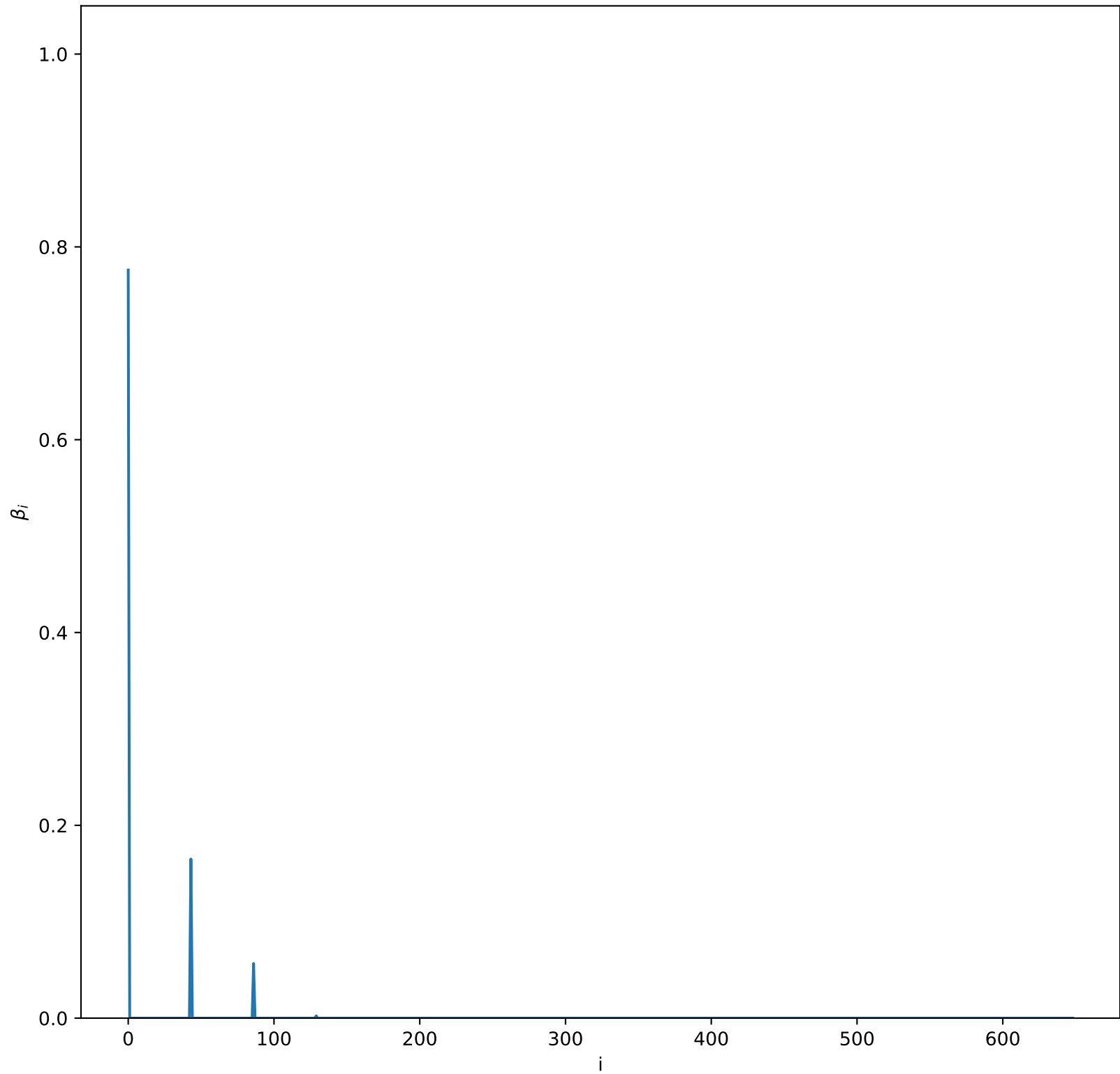
$\mu = 1.85$



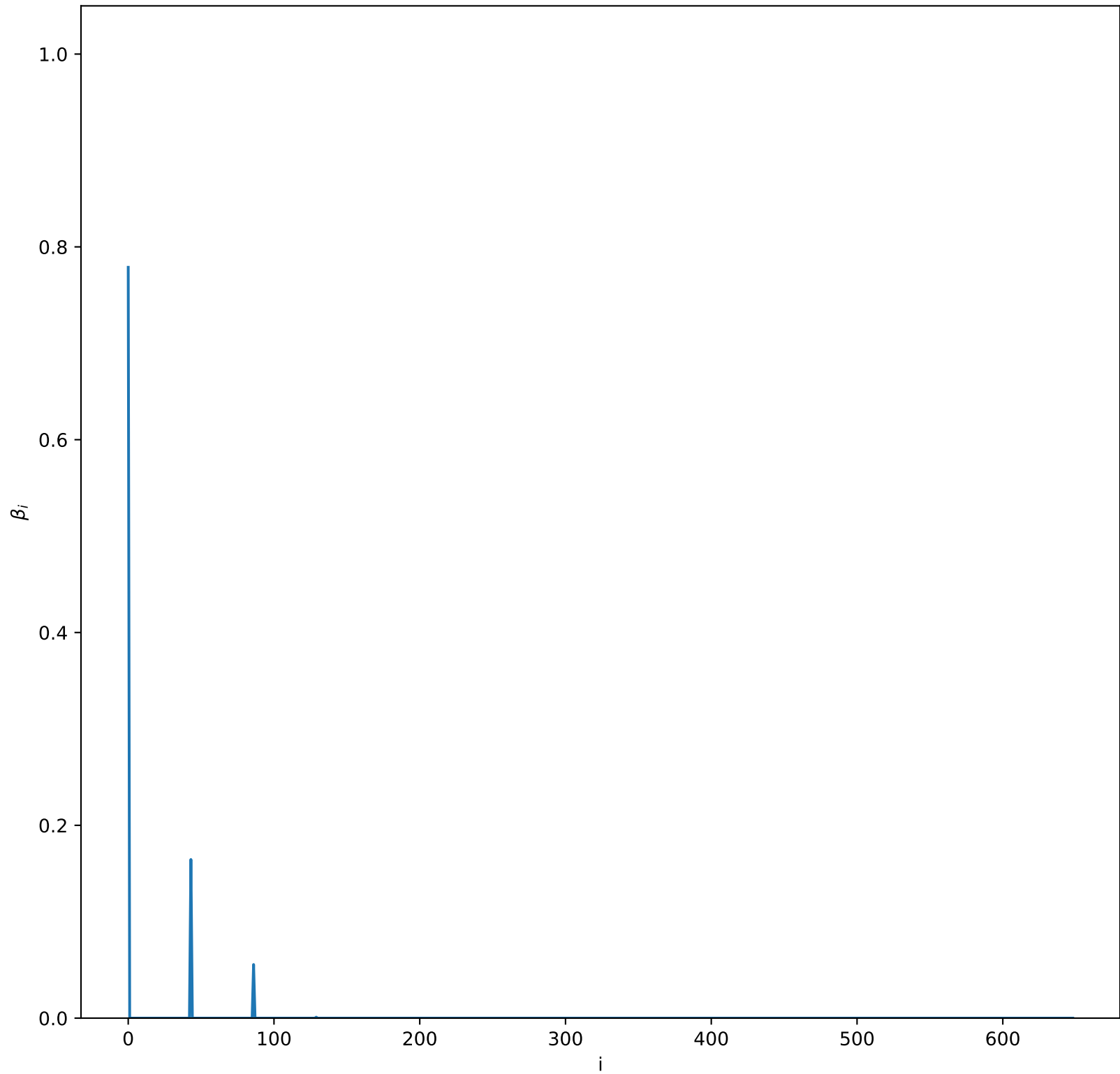
$\mu = 1.86$



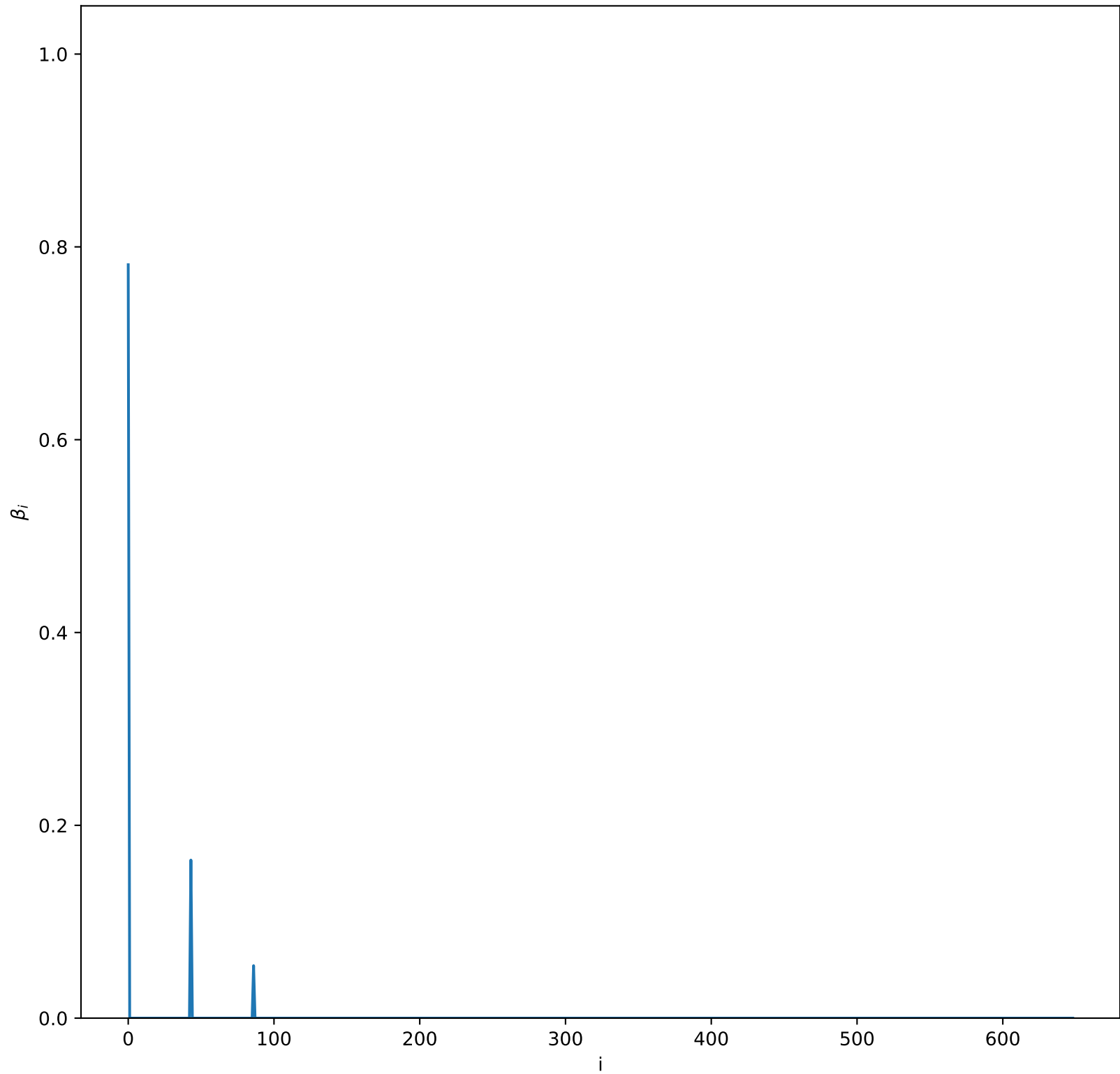
$\mu = 1.87$



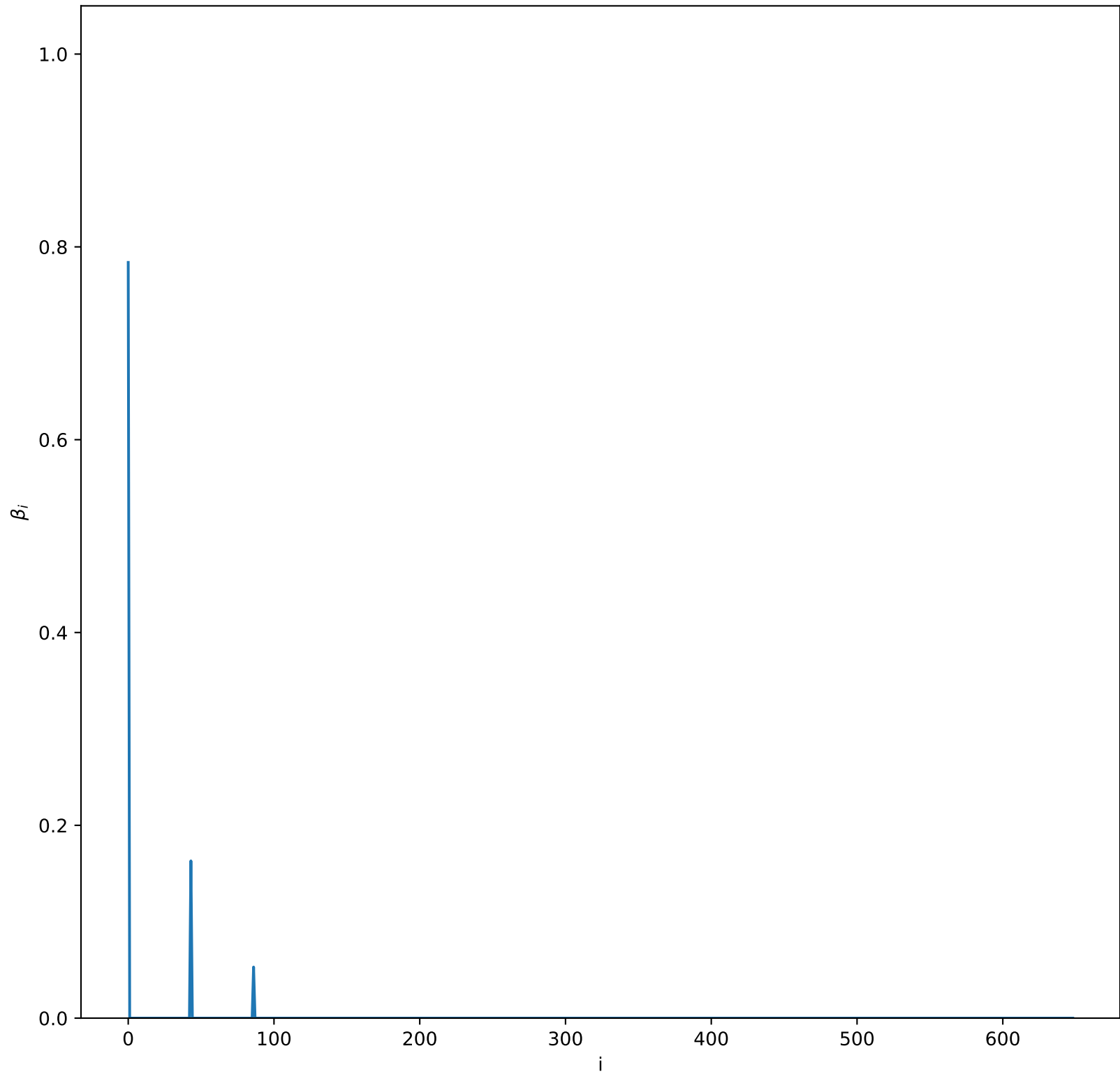
$\mu = 1.88$



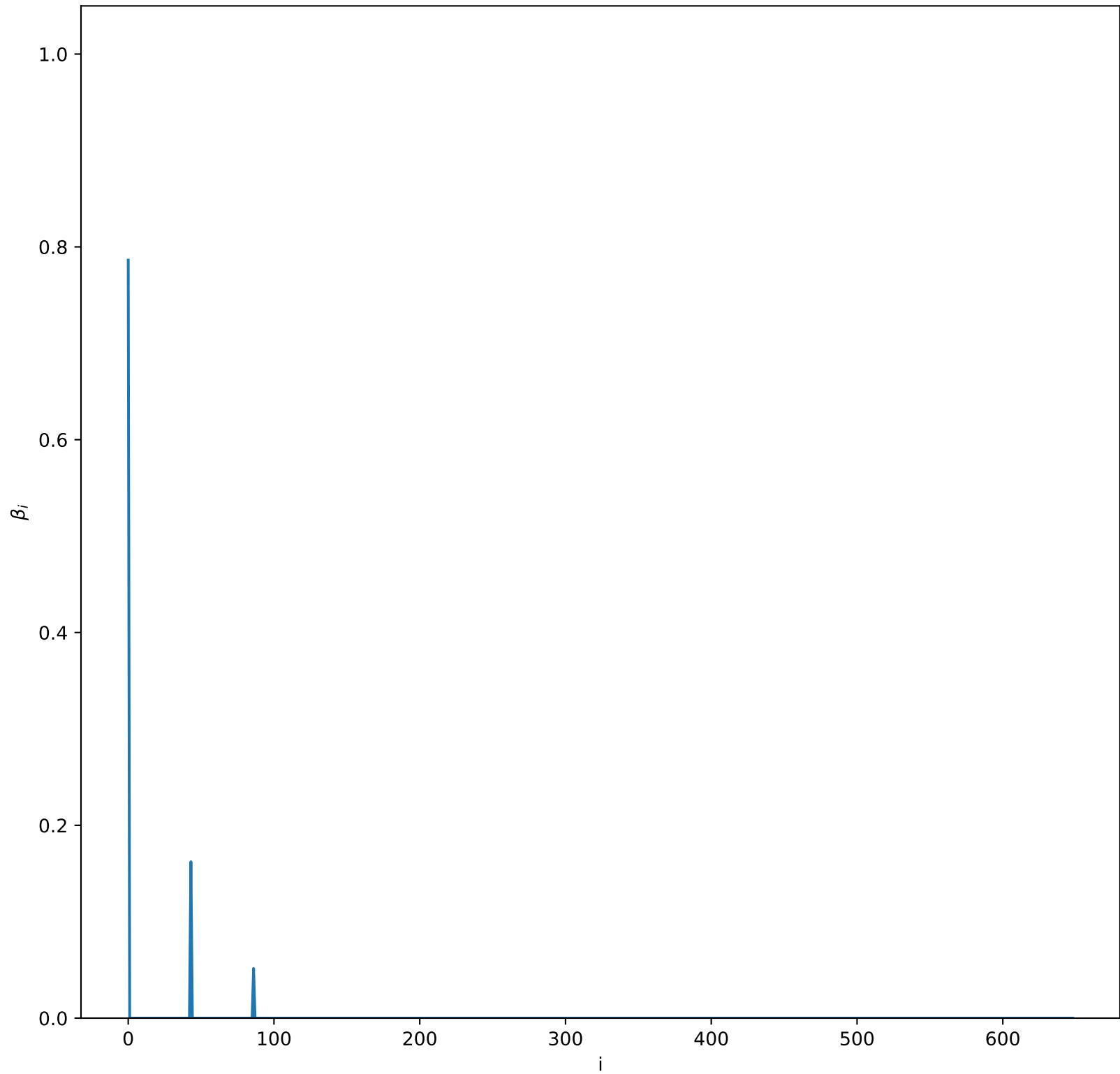
$\mu = 1.89$



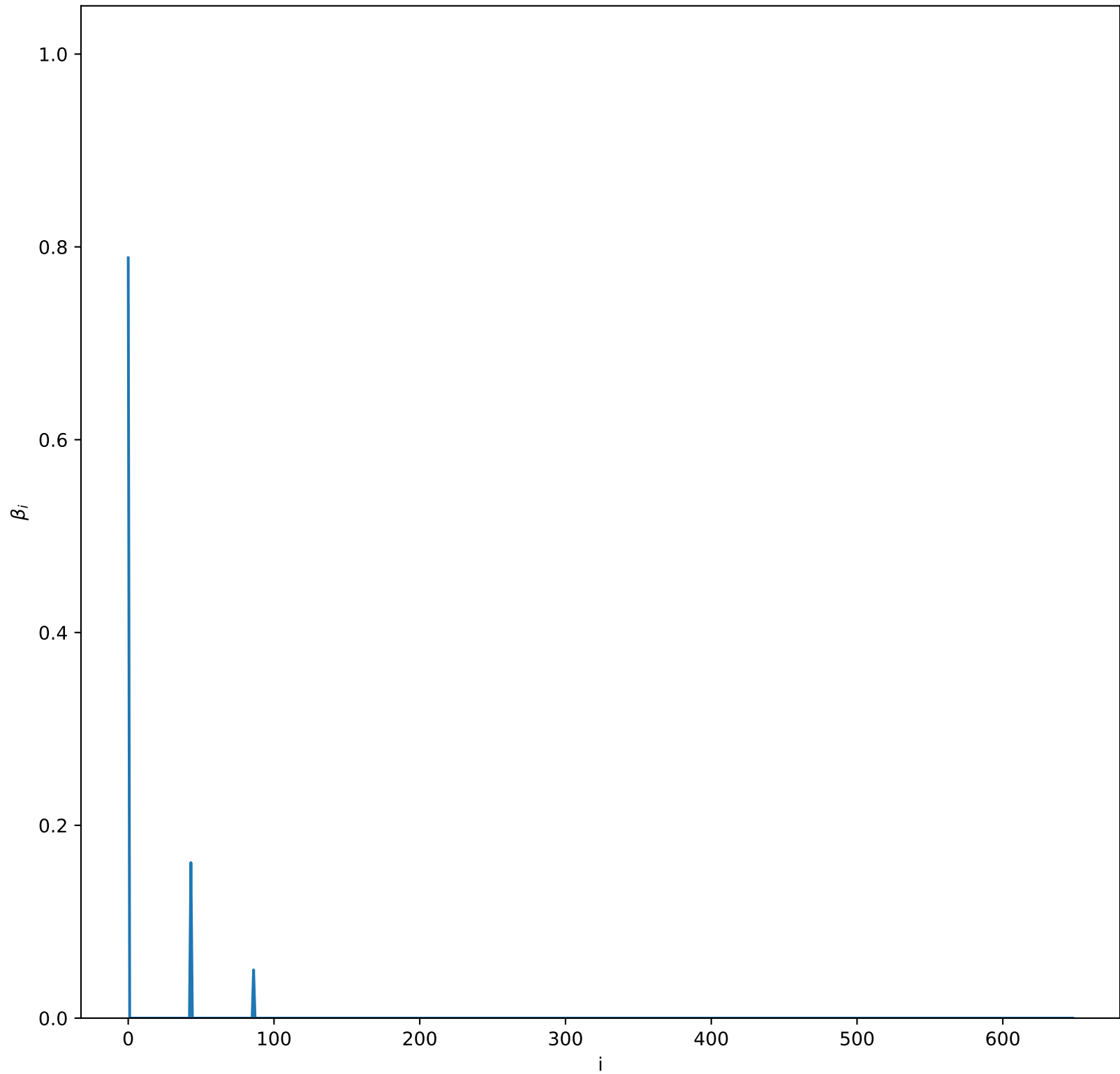
$\mu = 1.90$



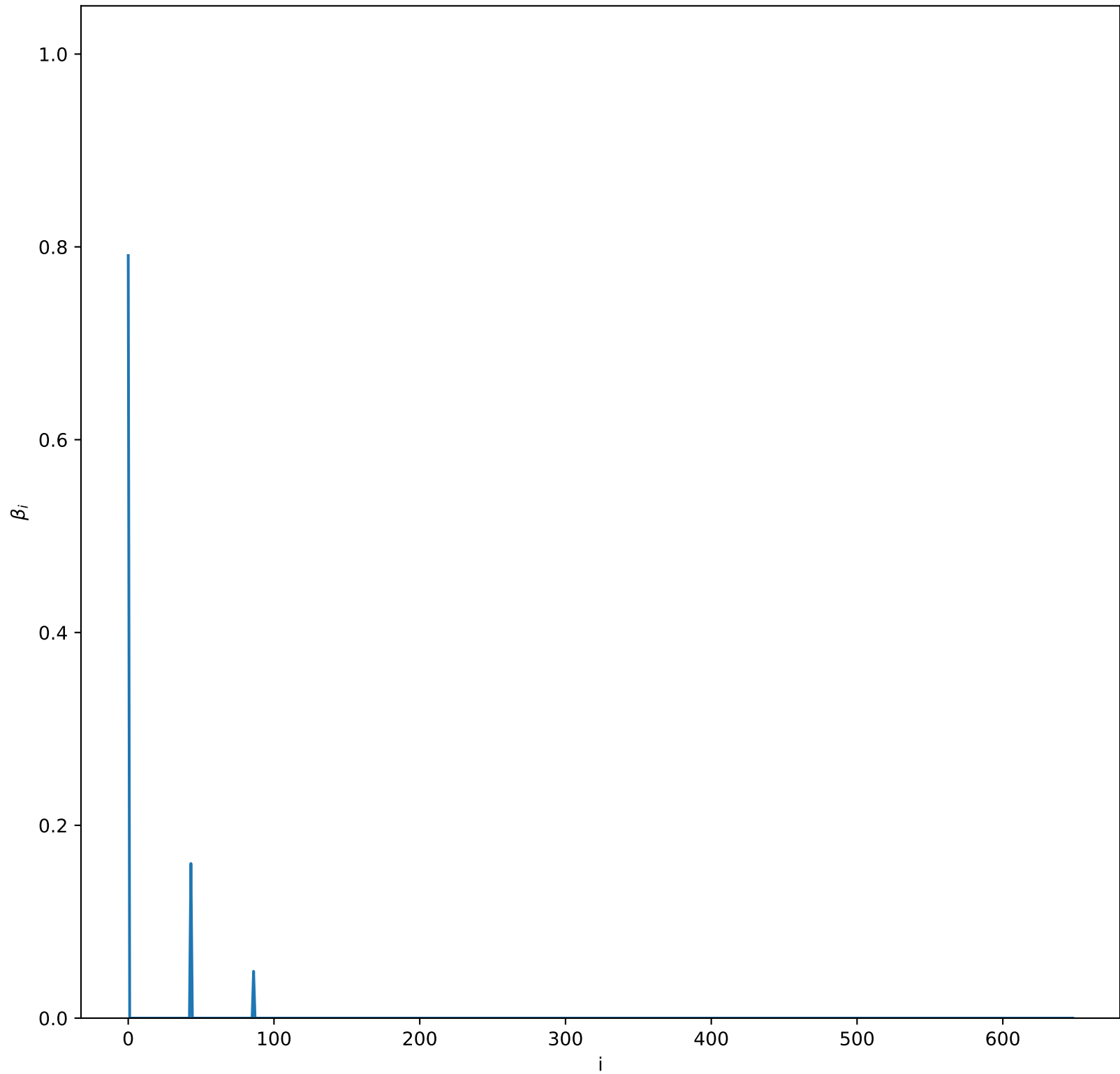
$\mu = 1.91$



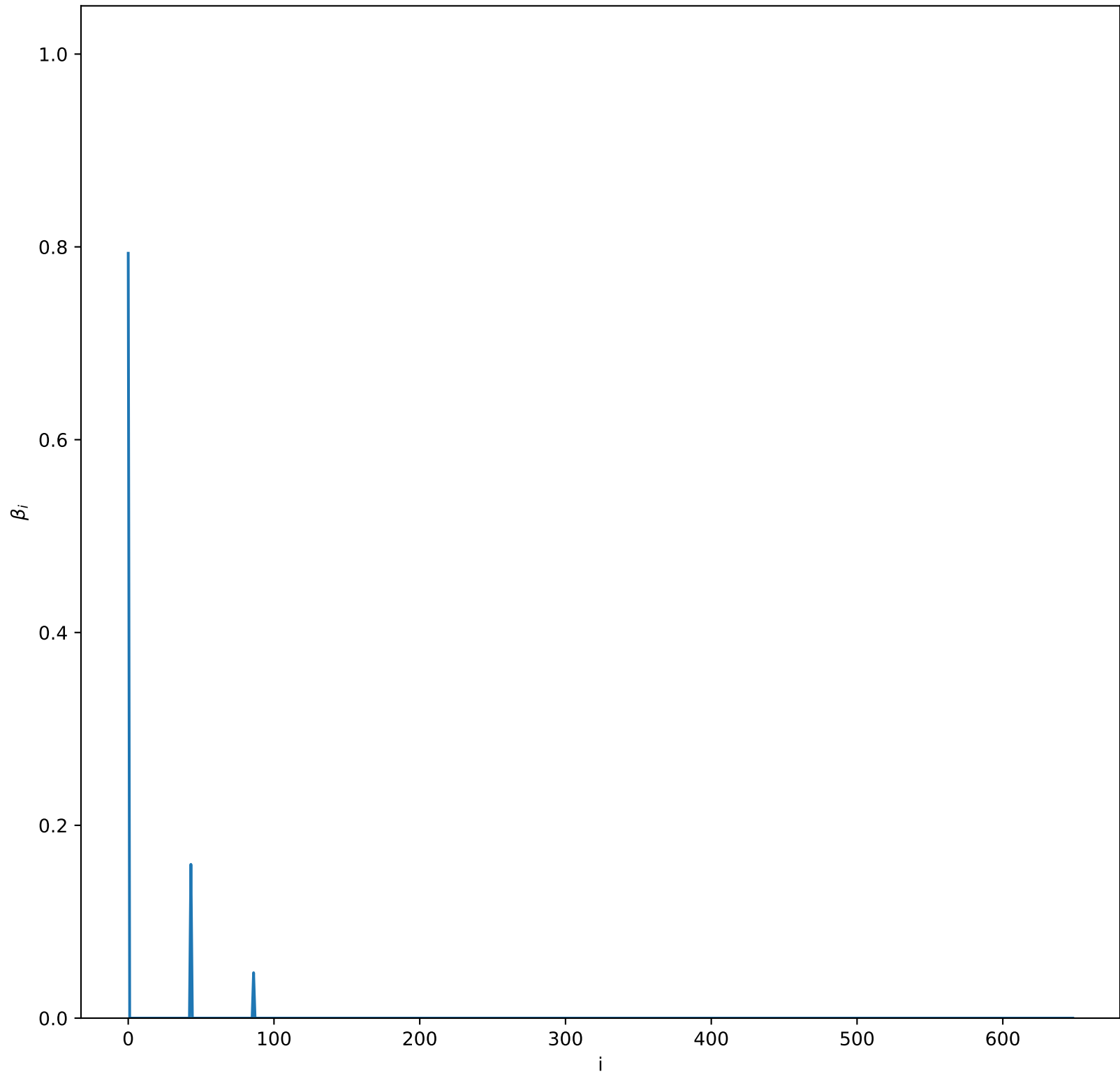
$\mu = 1.92$



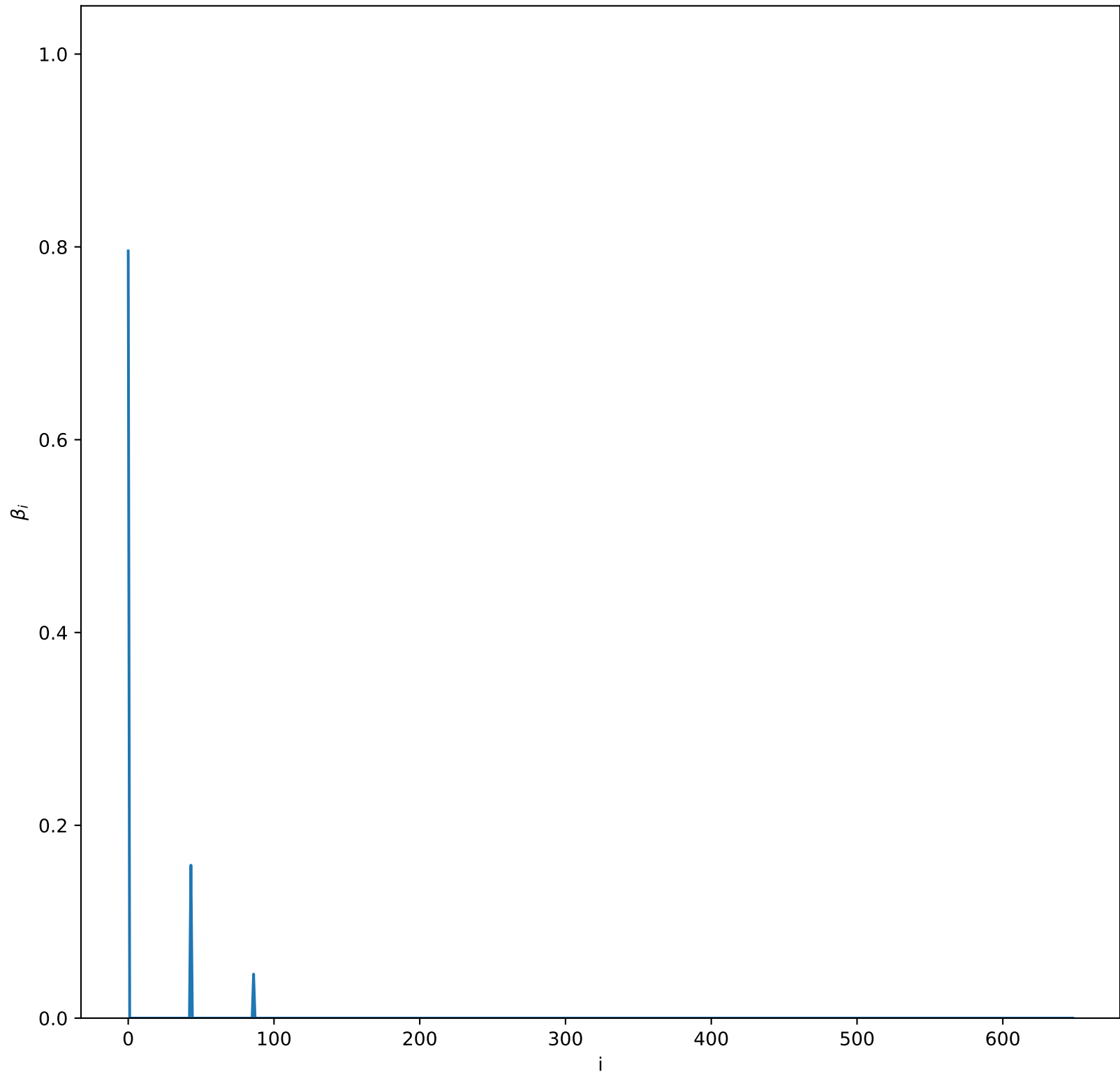
$\mu = 1.93$



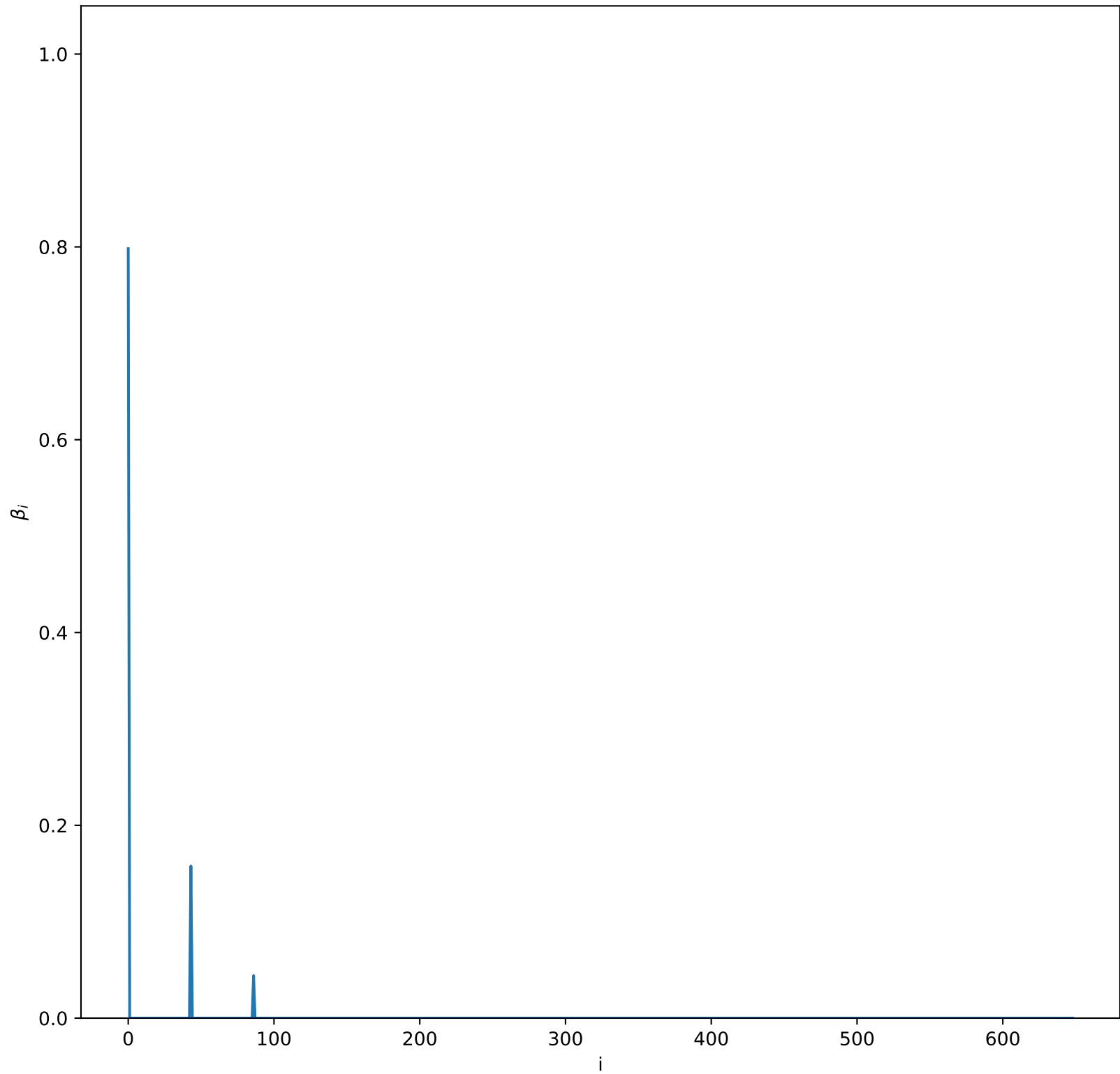
$\mu = 1.94$



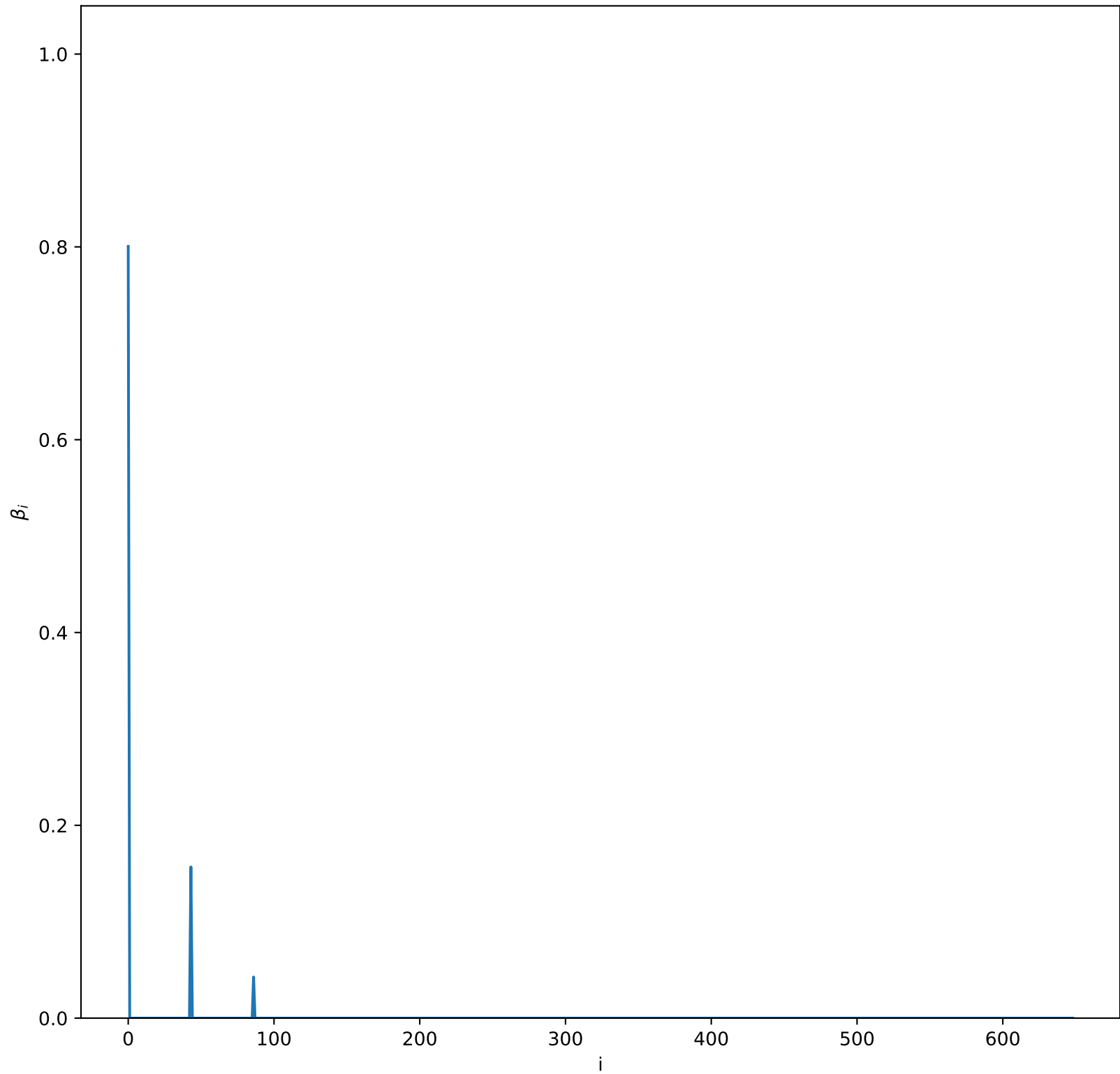
$\mu = 1.95$



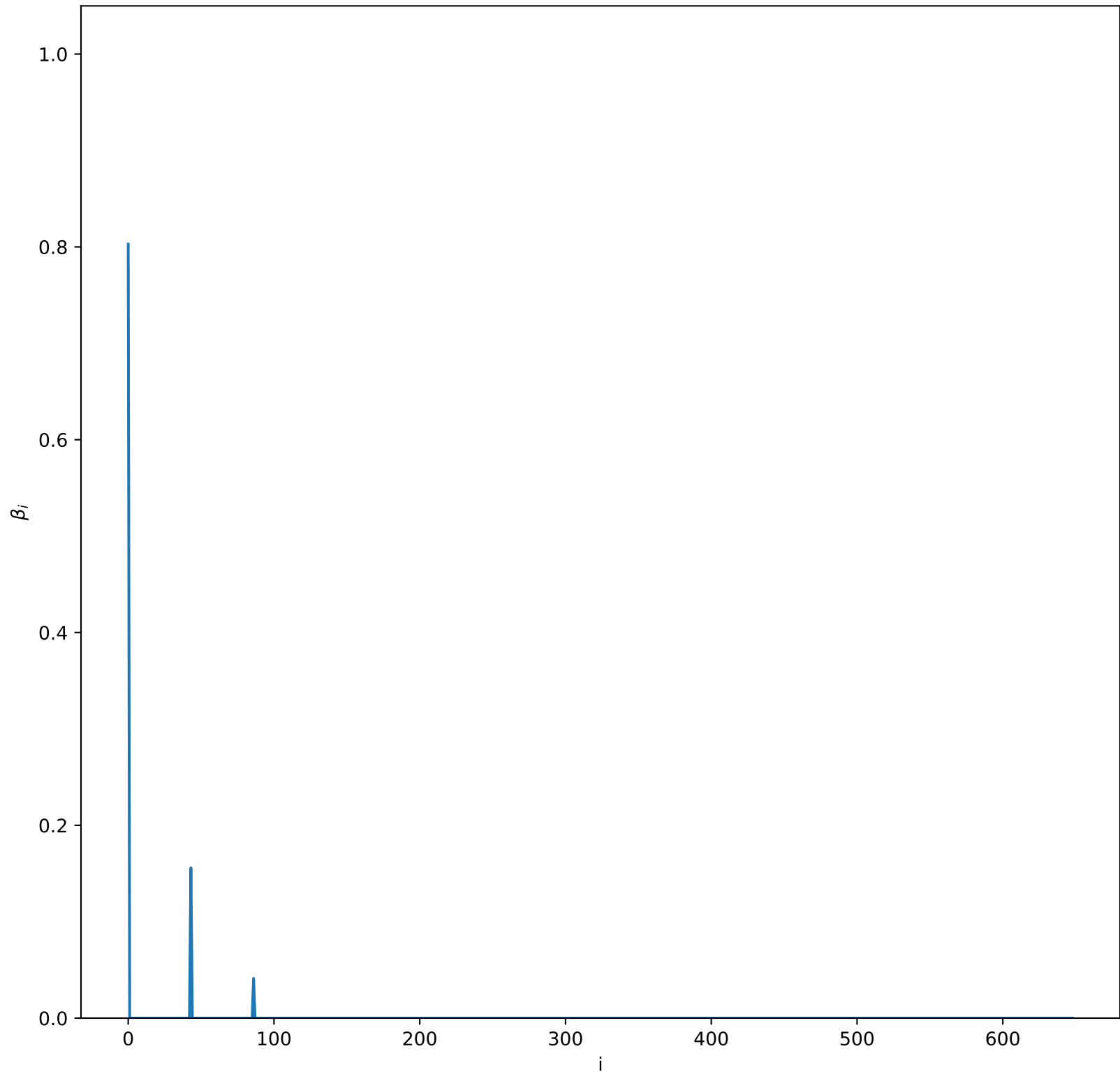
$\mu = 1.96$



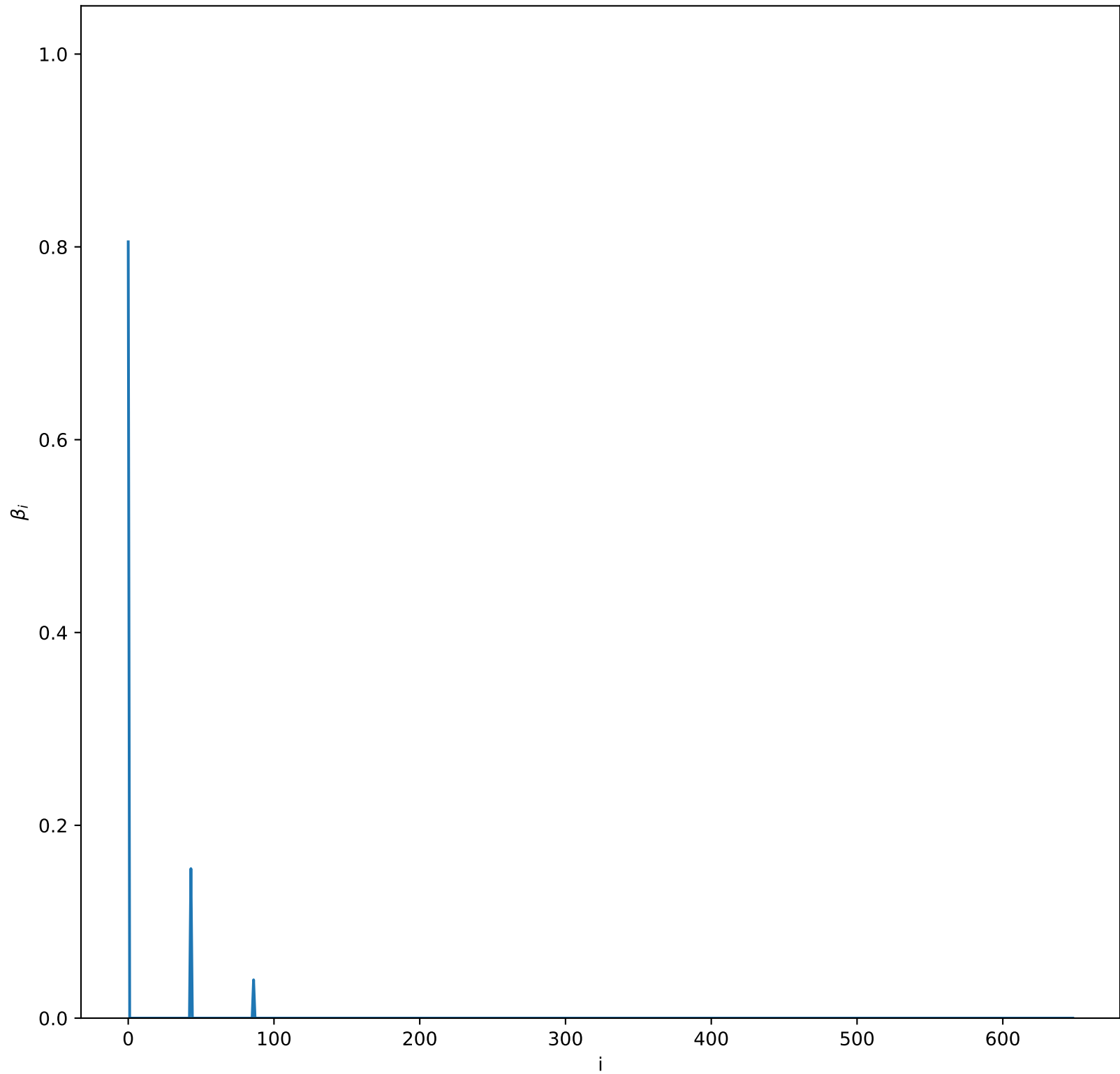
$\mu = 1.97$



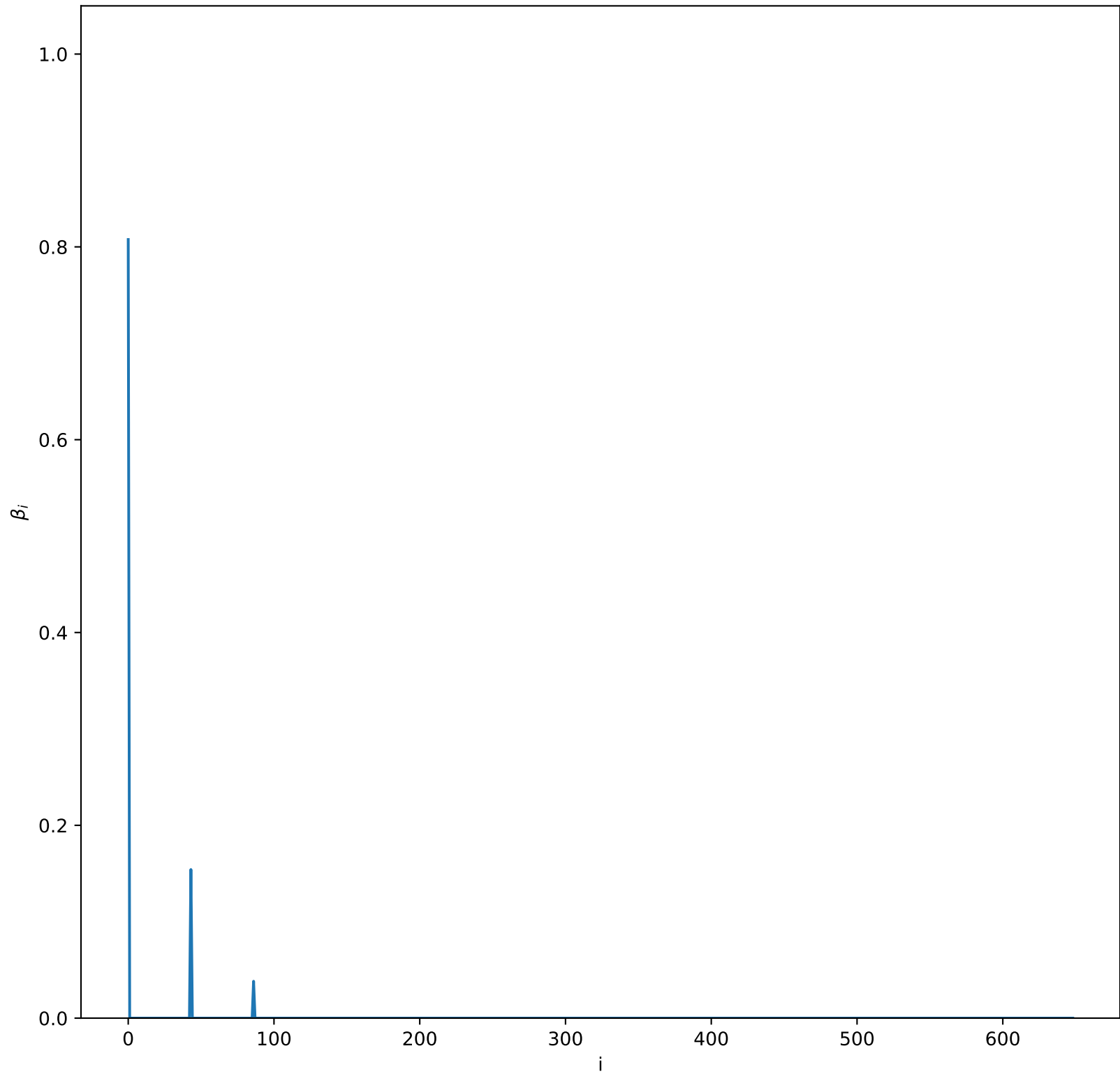
$\mu = 1.98$



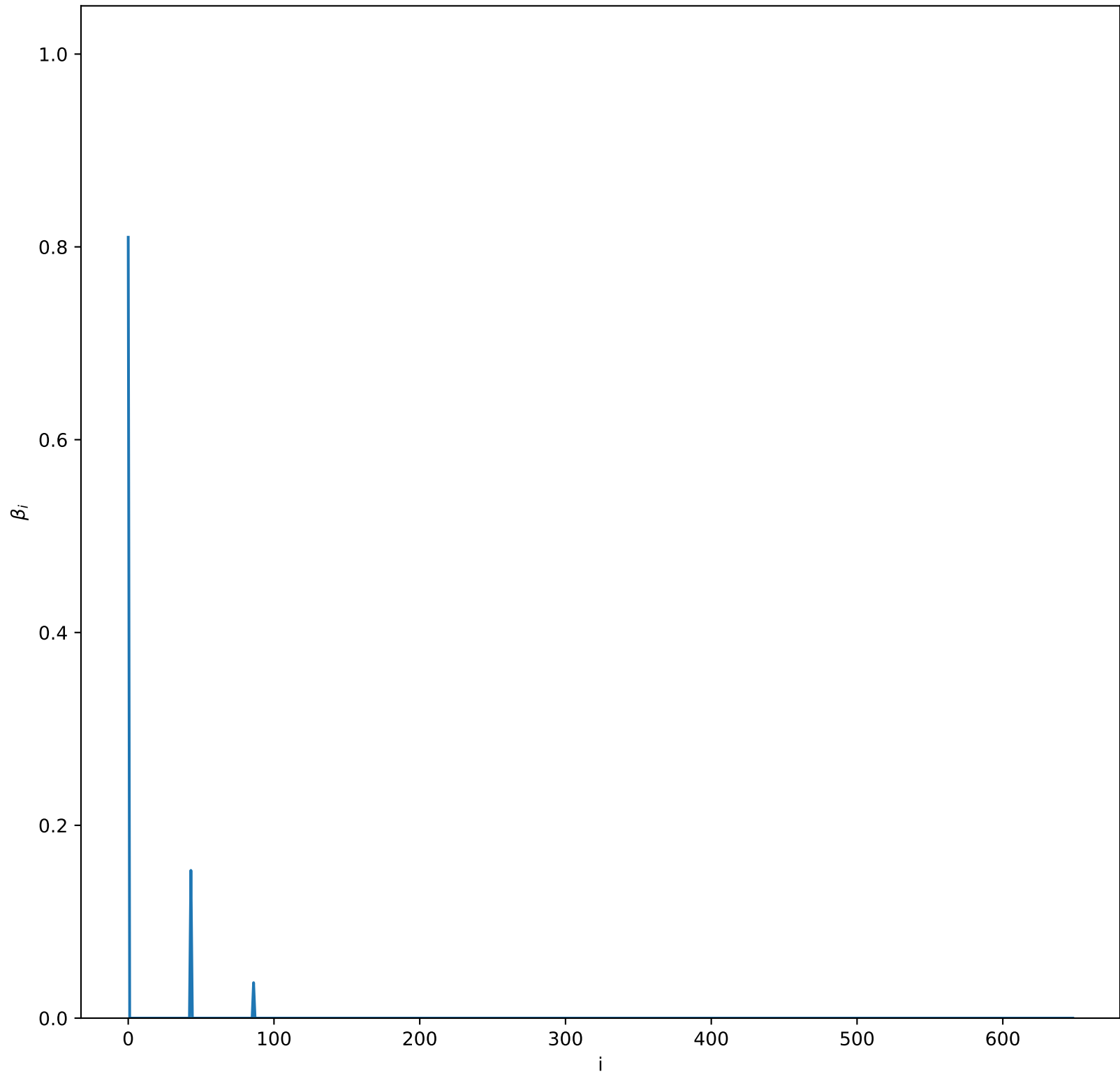
$\mu = 1.99$



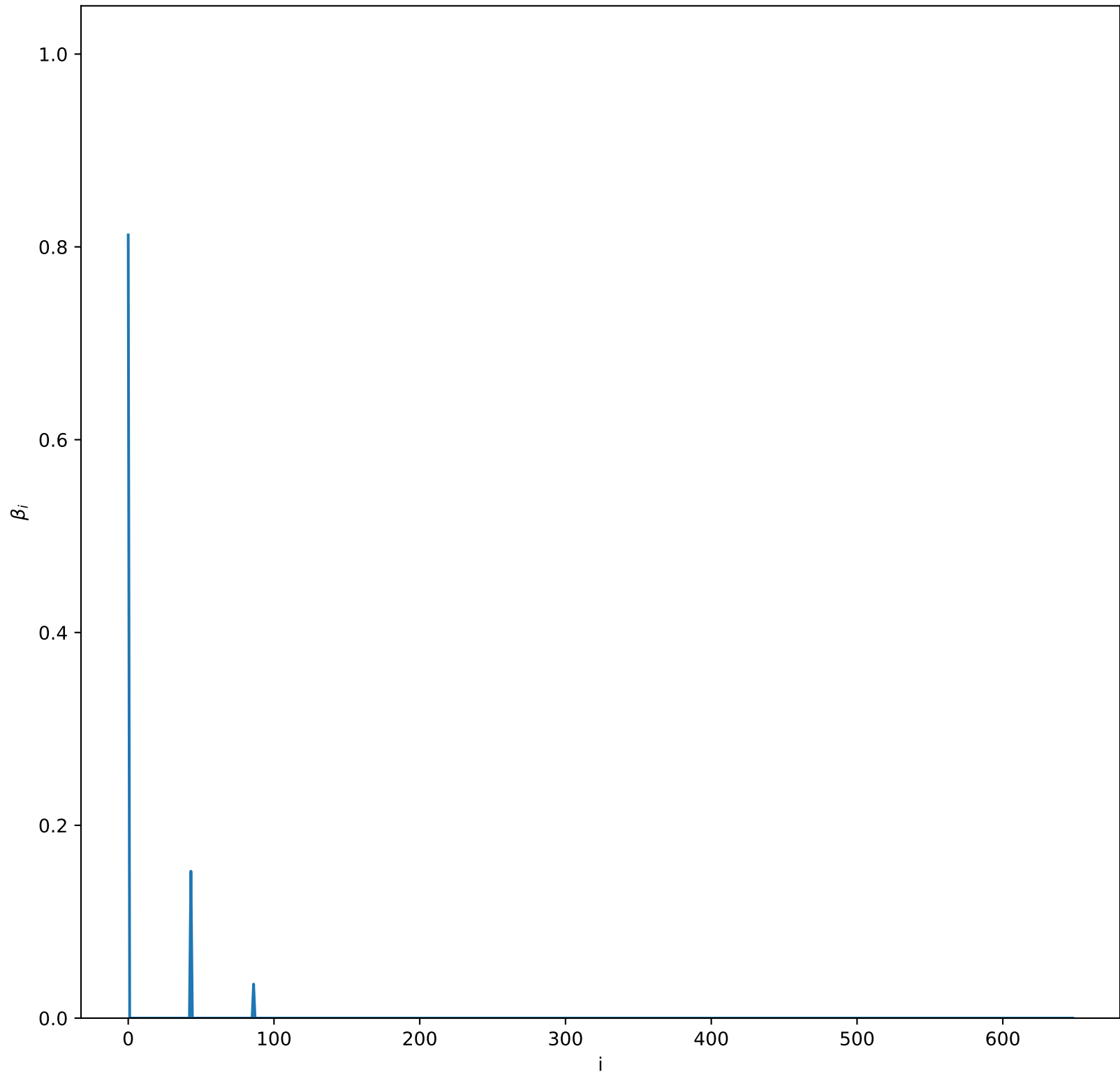
$\mu = 2.00$



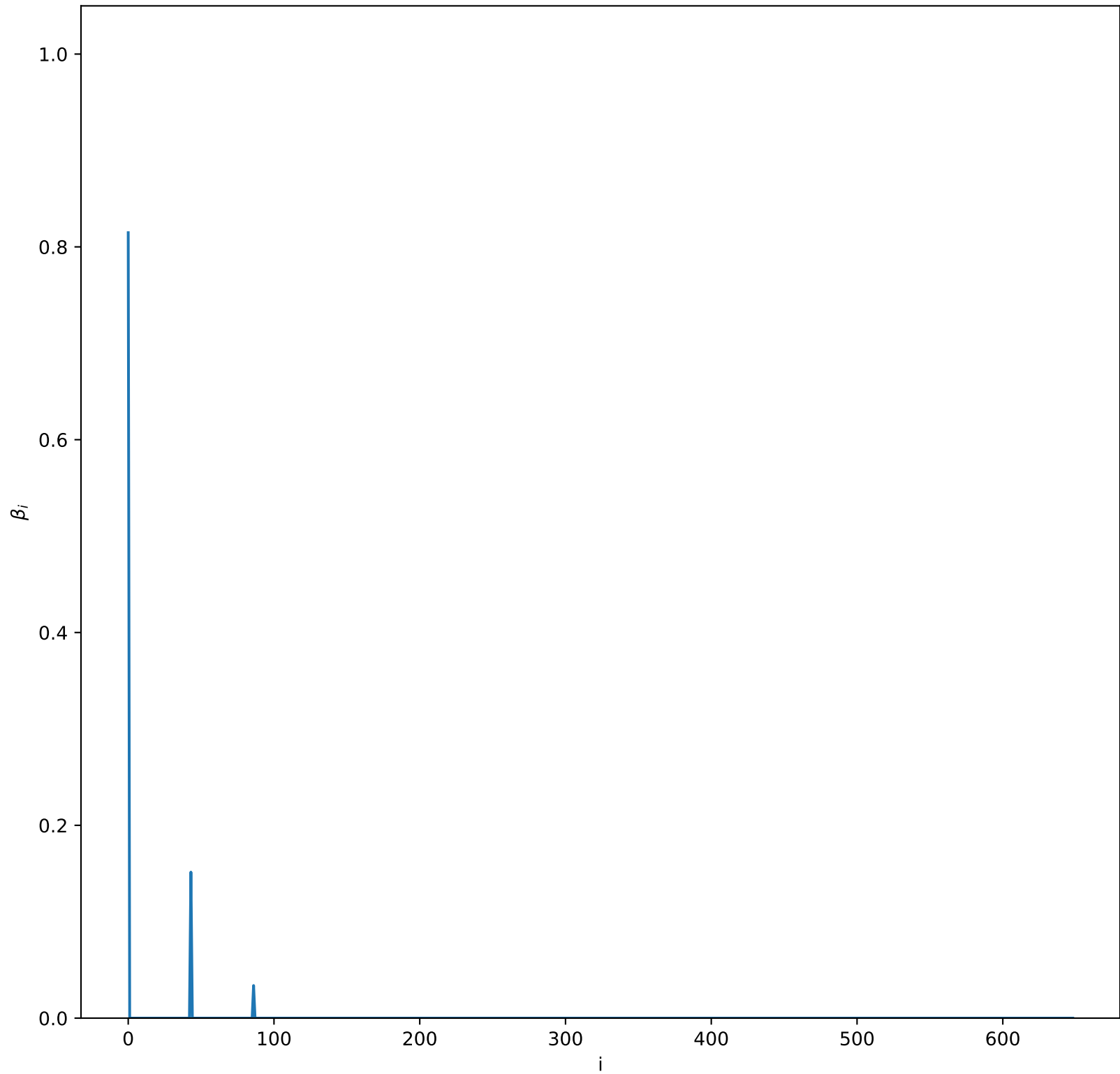
$\mu = 2.01$



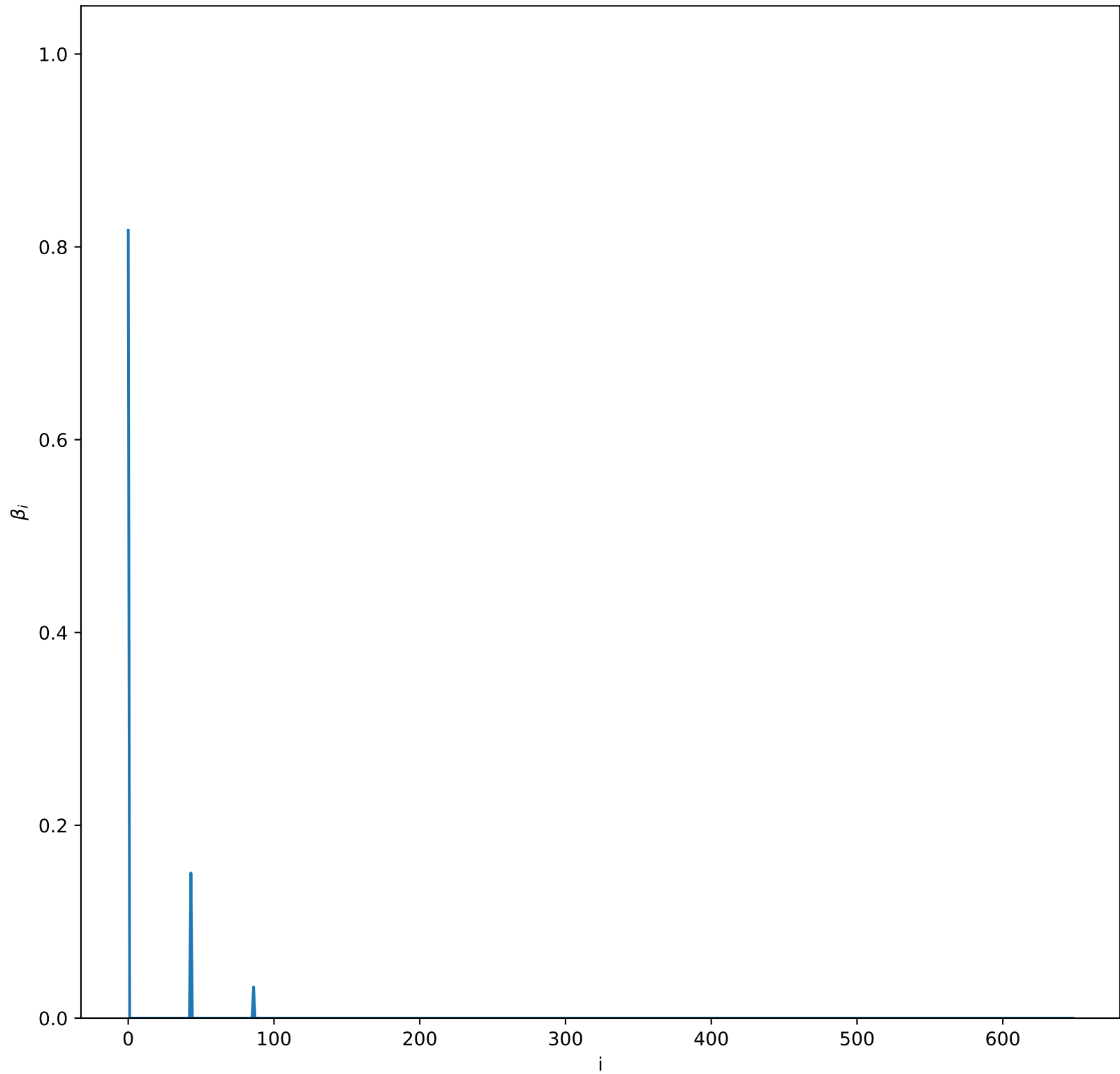
$\mu = 2.02$



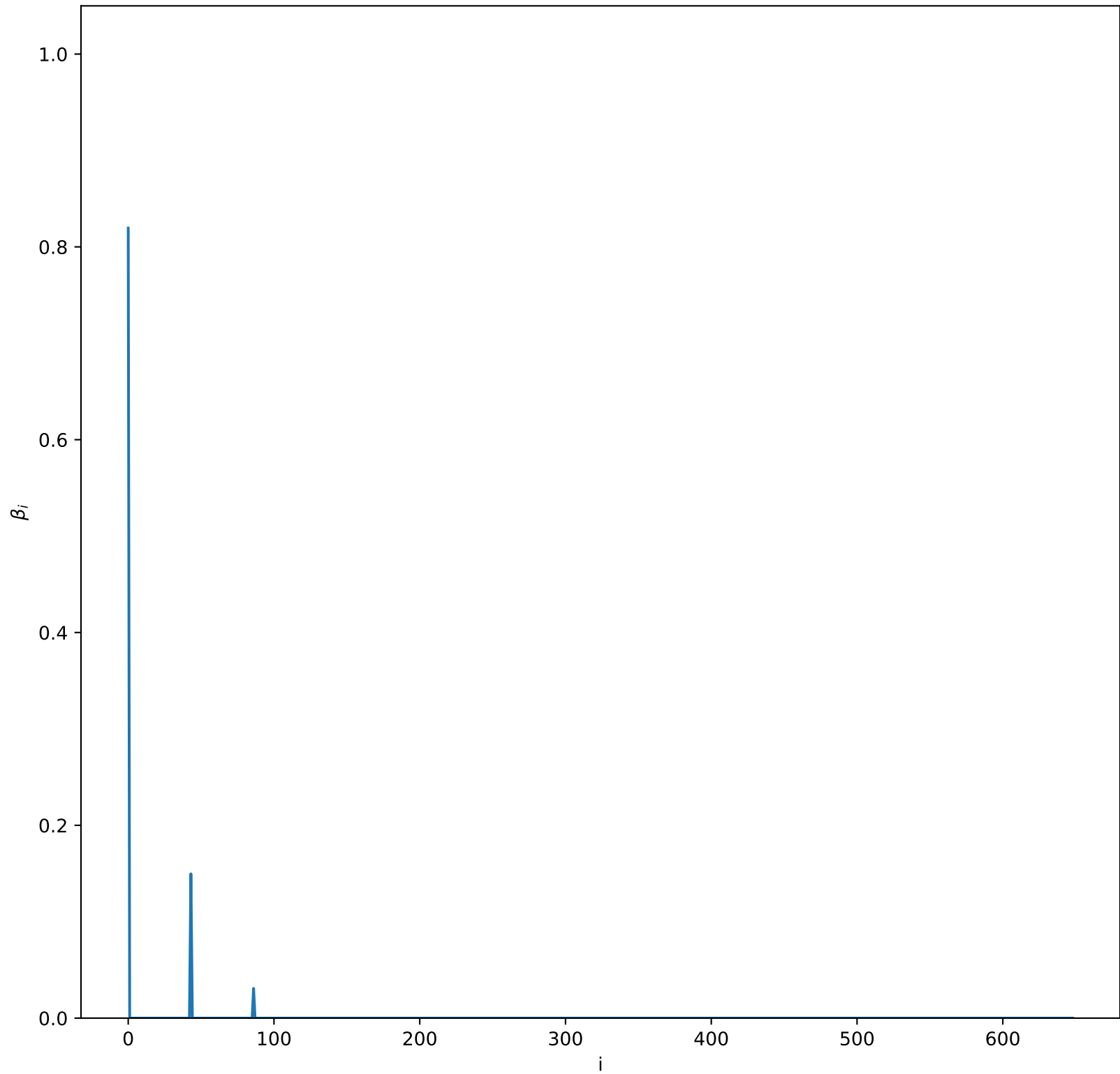
$\mu = 2.03$



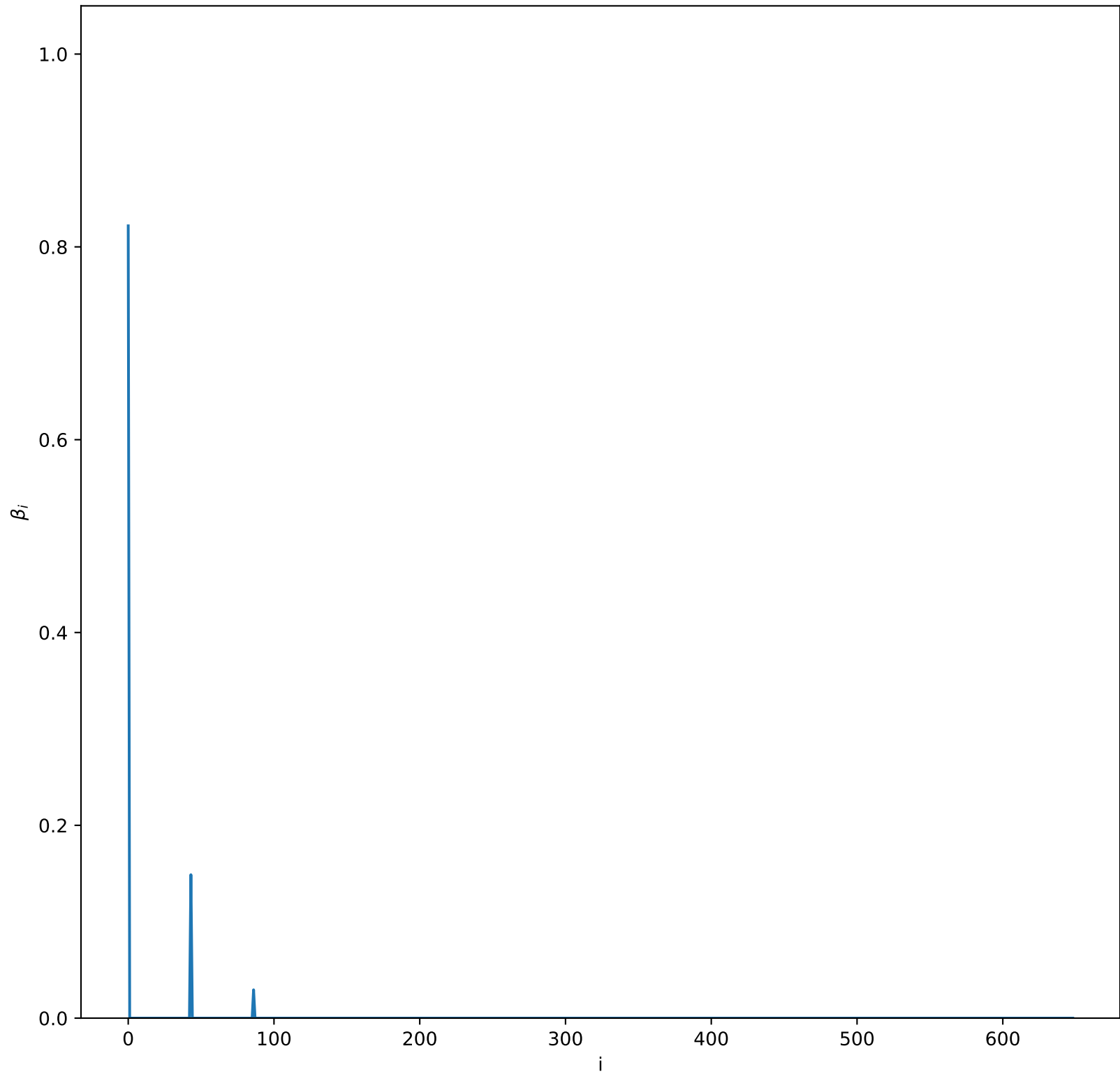
$\mu = 2.04$



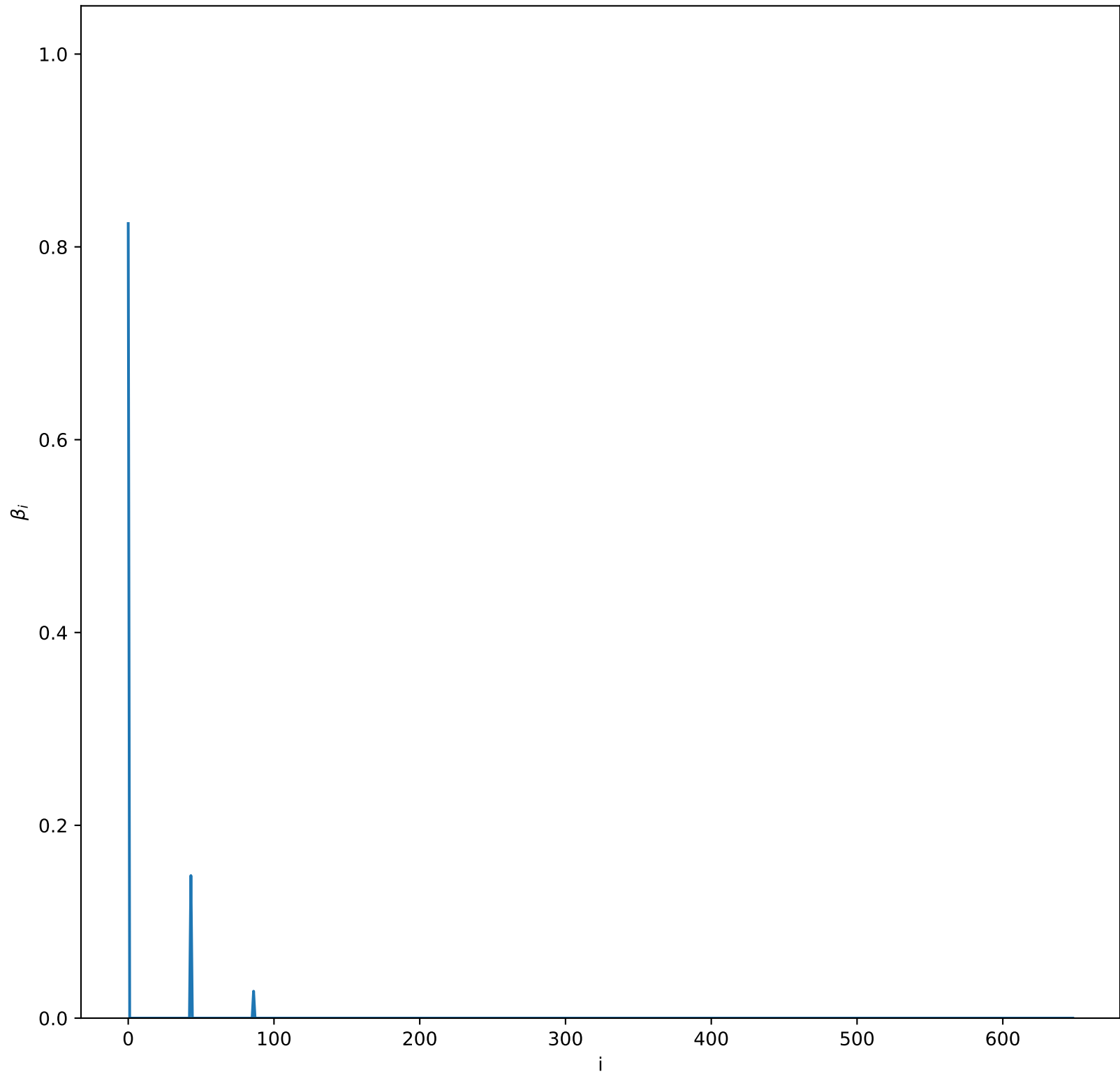
$\mu = 2.05$



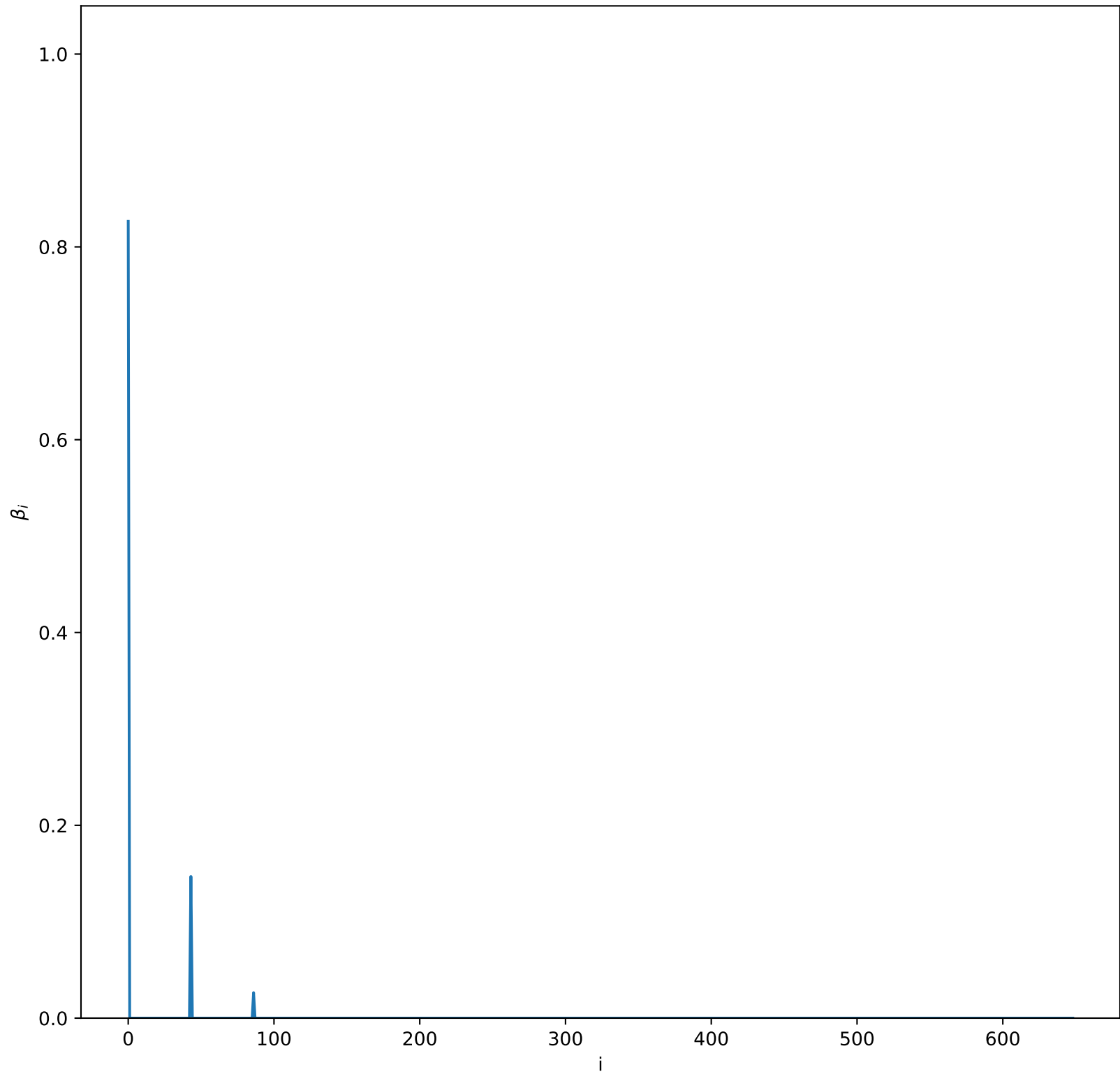
$\mu = 2.06$



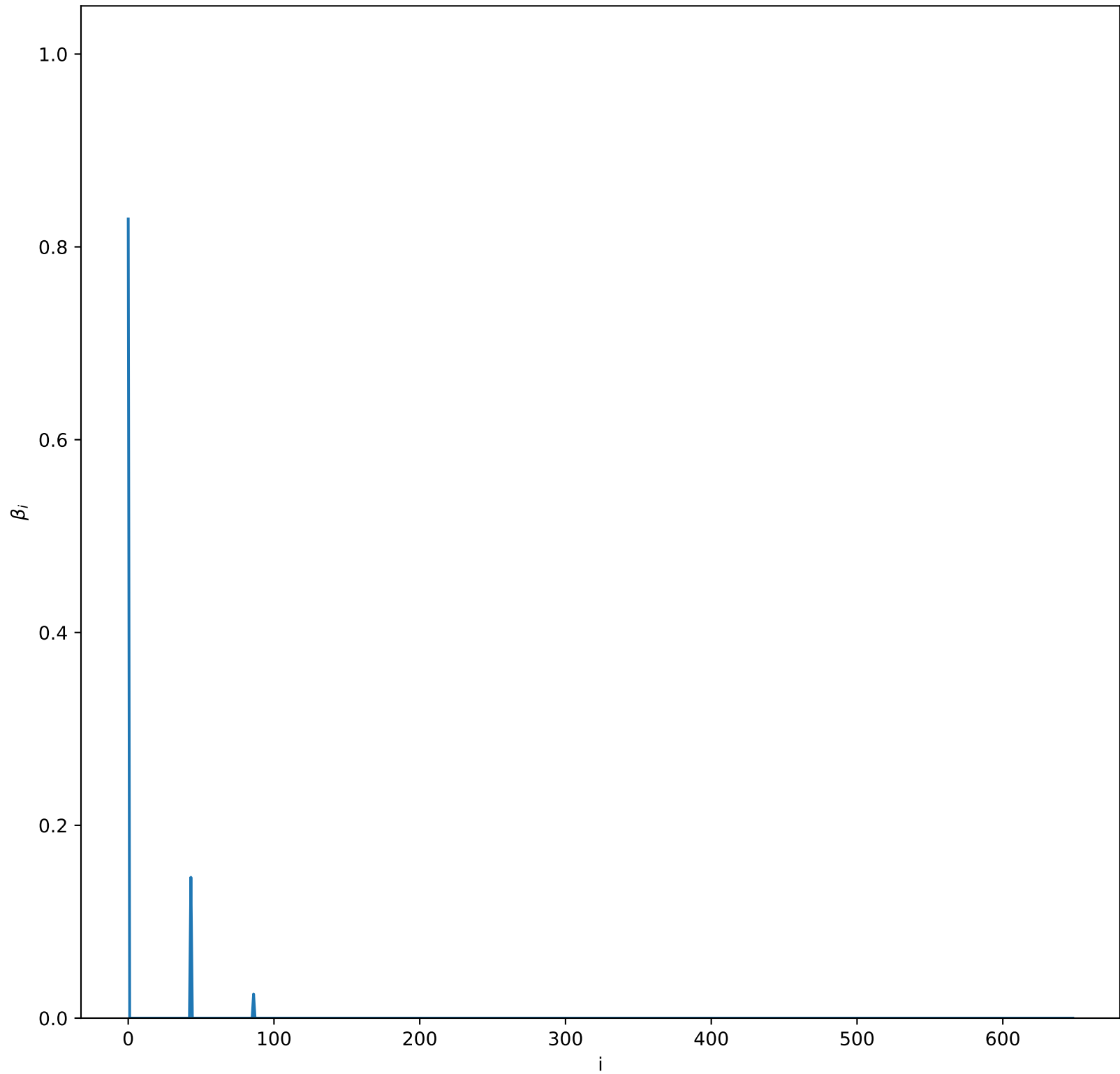
$\mu = 2.07$



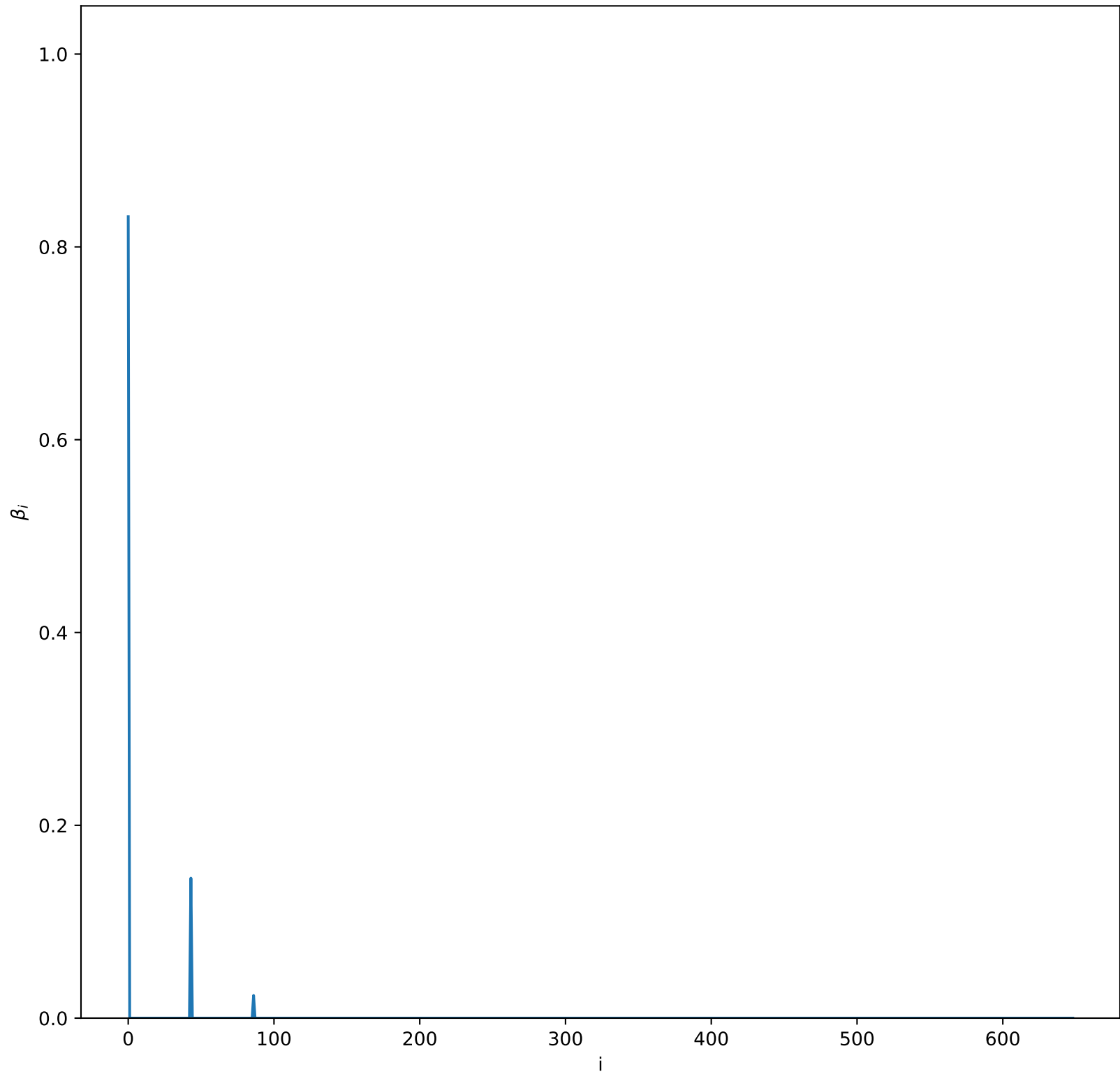
$\mu = 2.08$



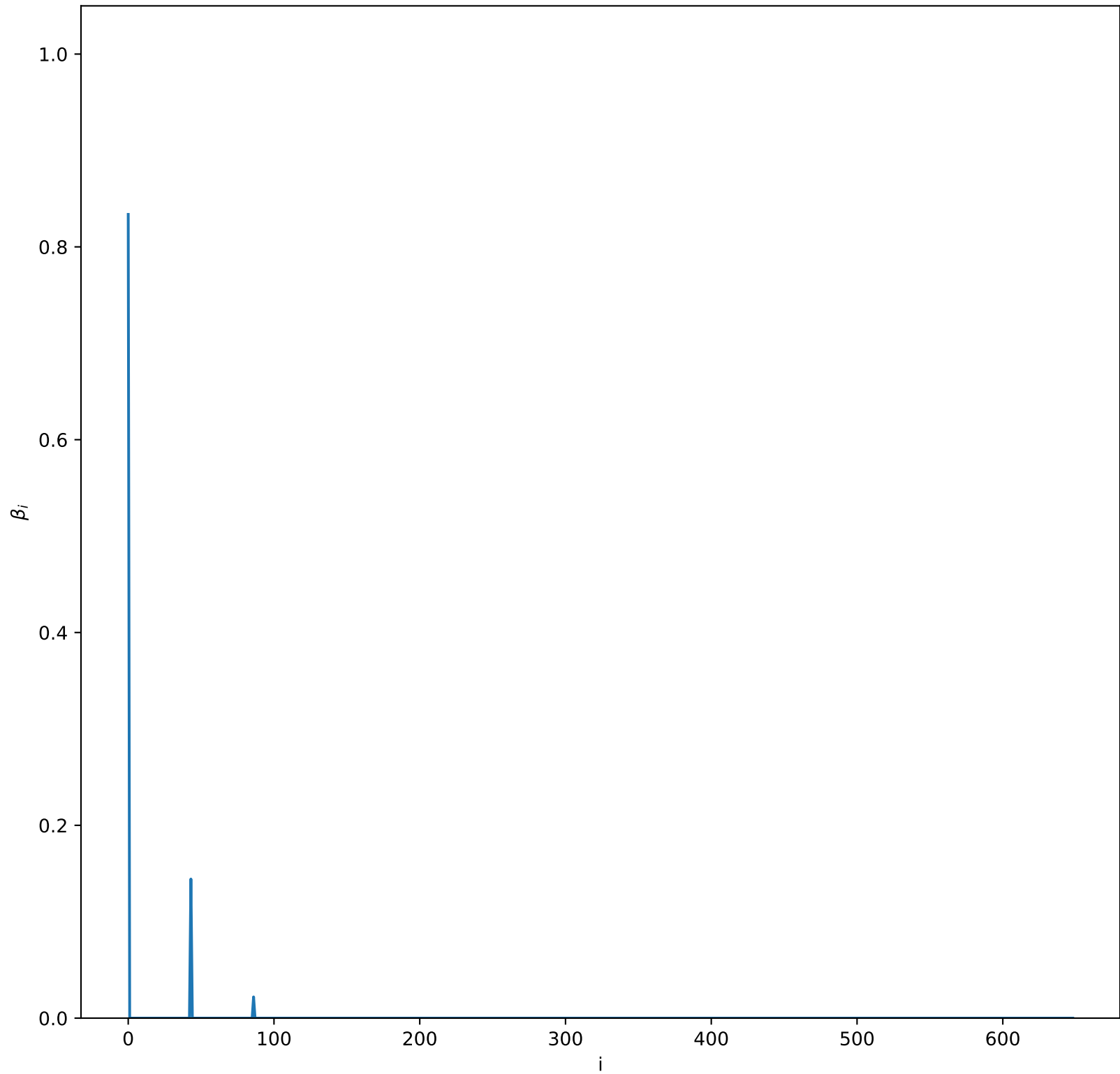
$\mu = 2.09$



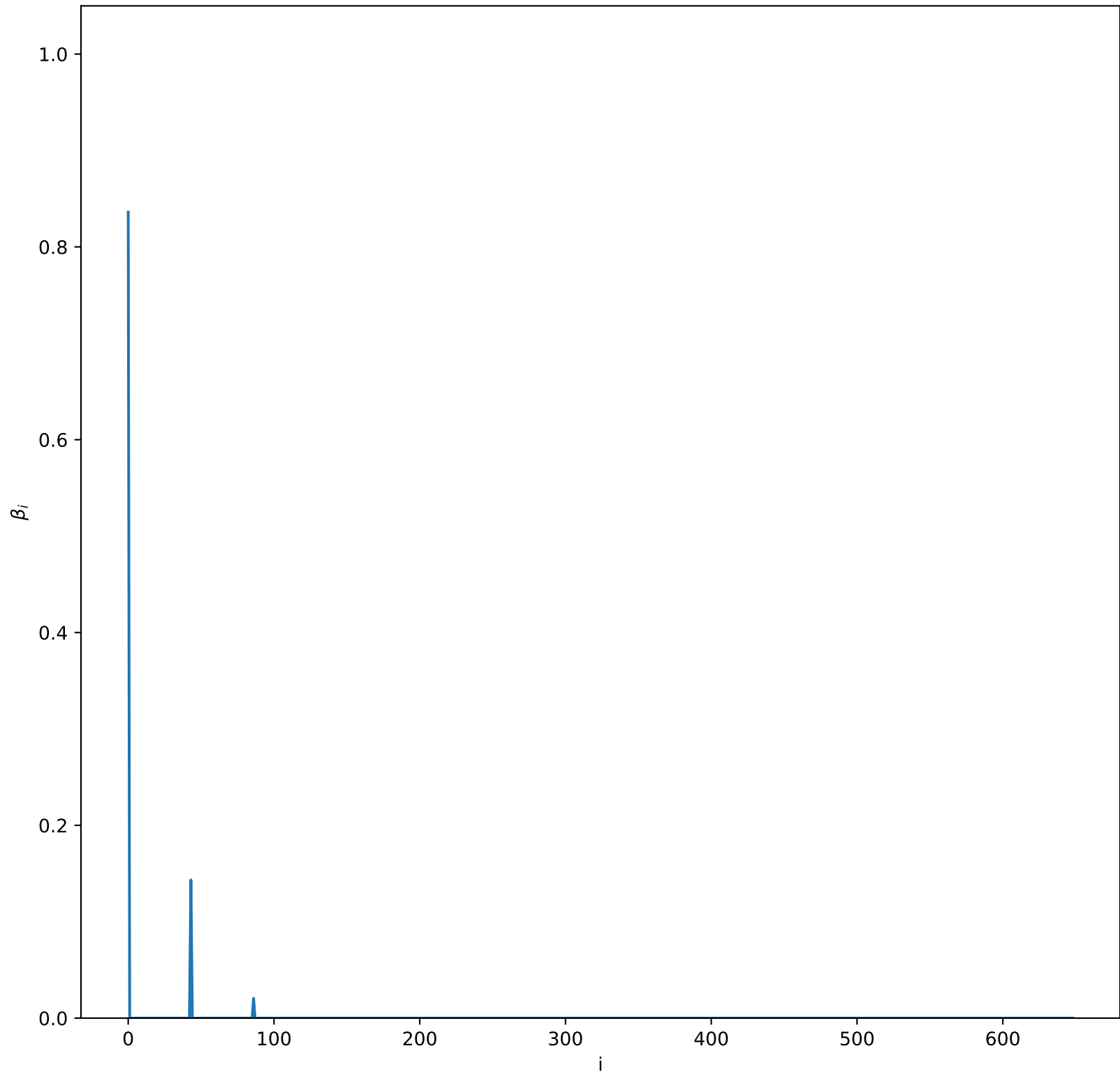
$\mu = 2.10$



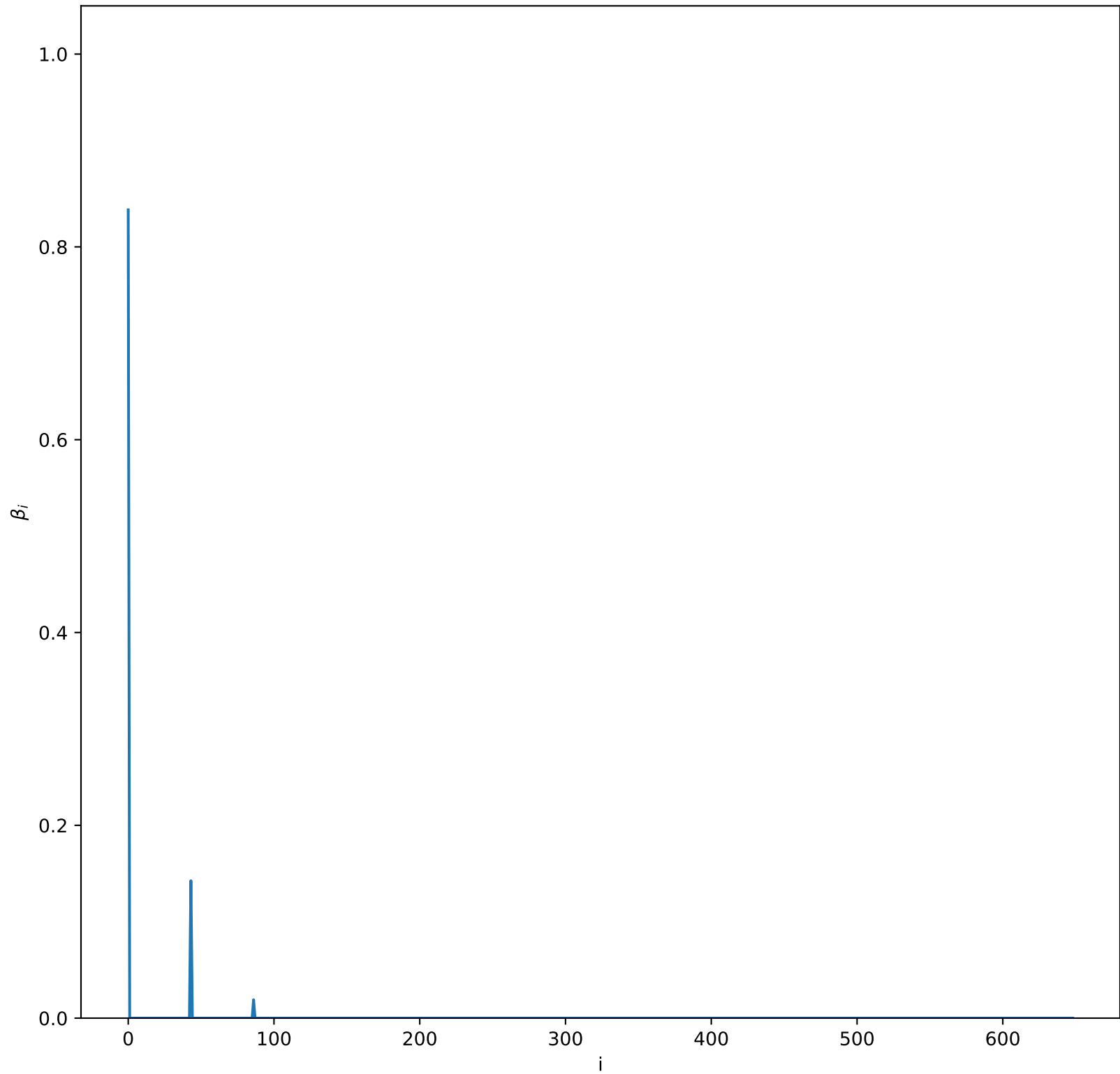
$\mu = 2.11$



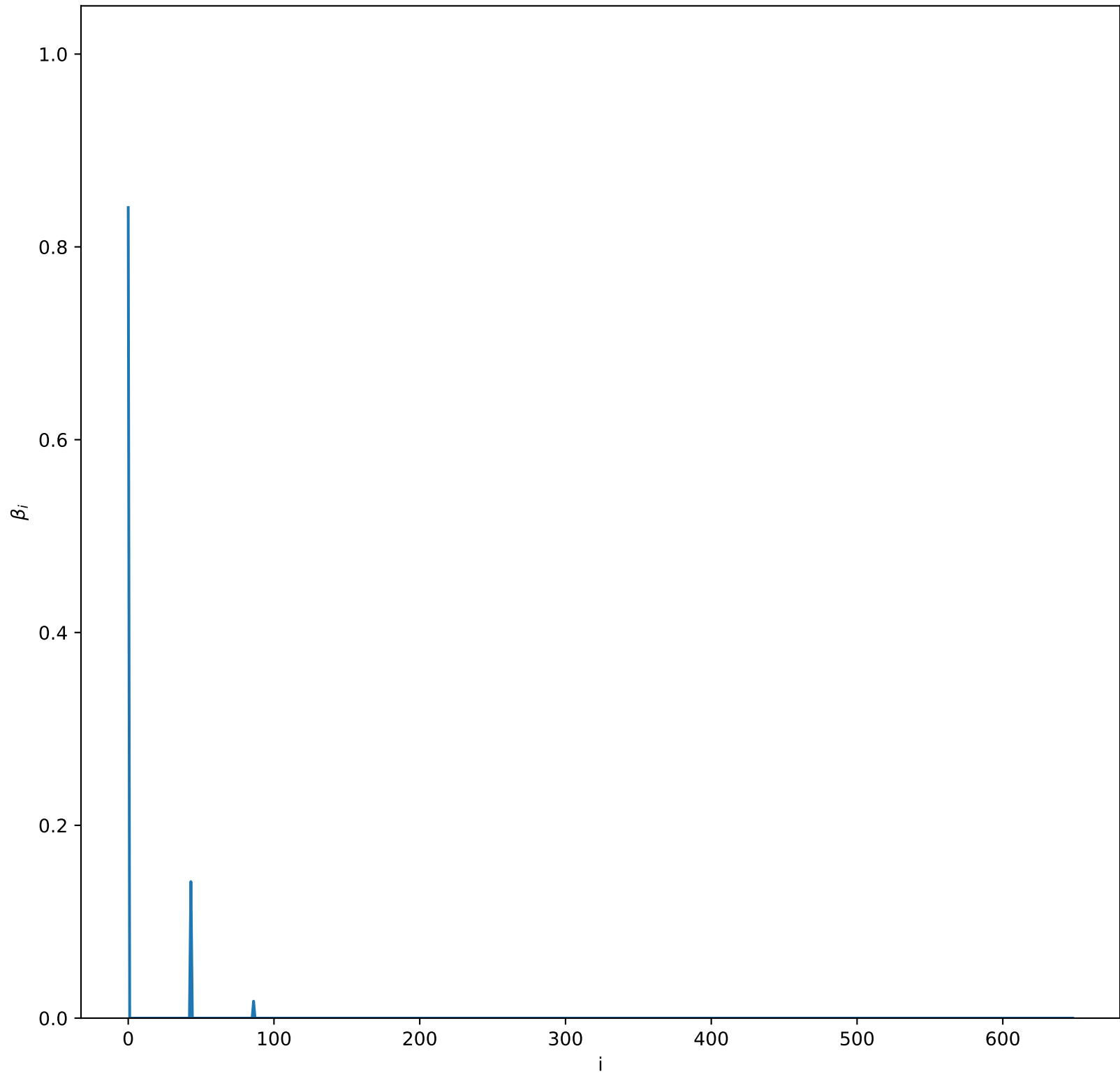
$\mu = 2.12$



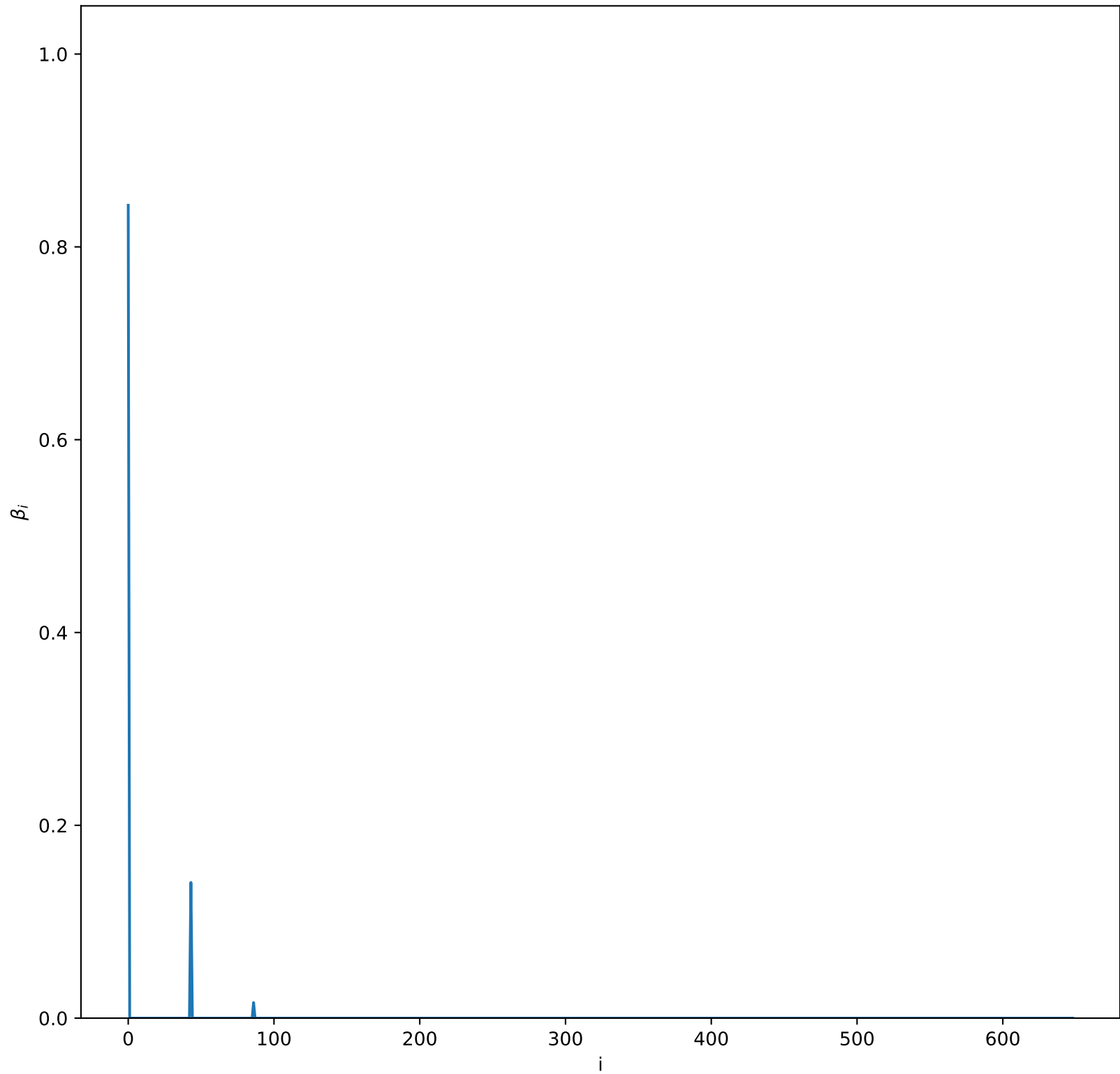
$\mu = 2.13$



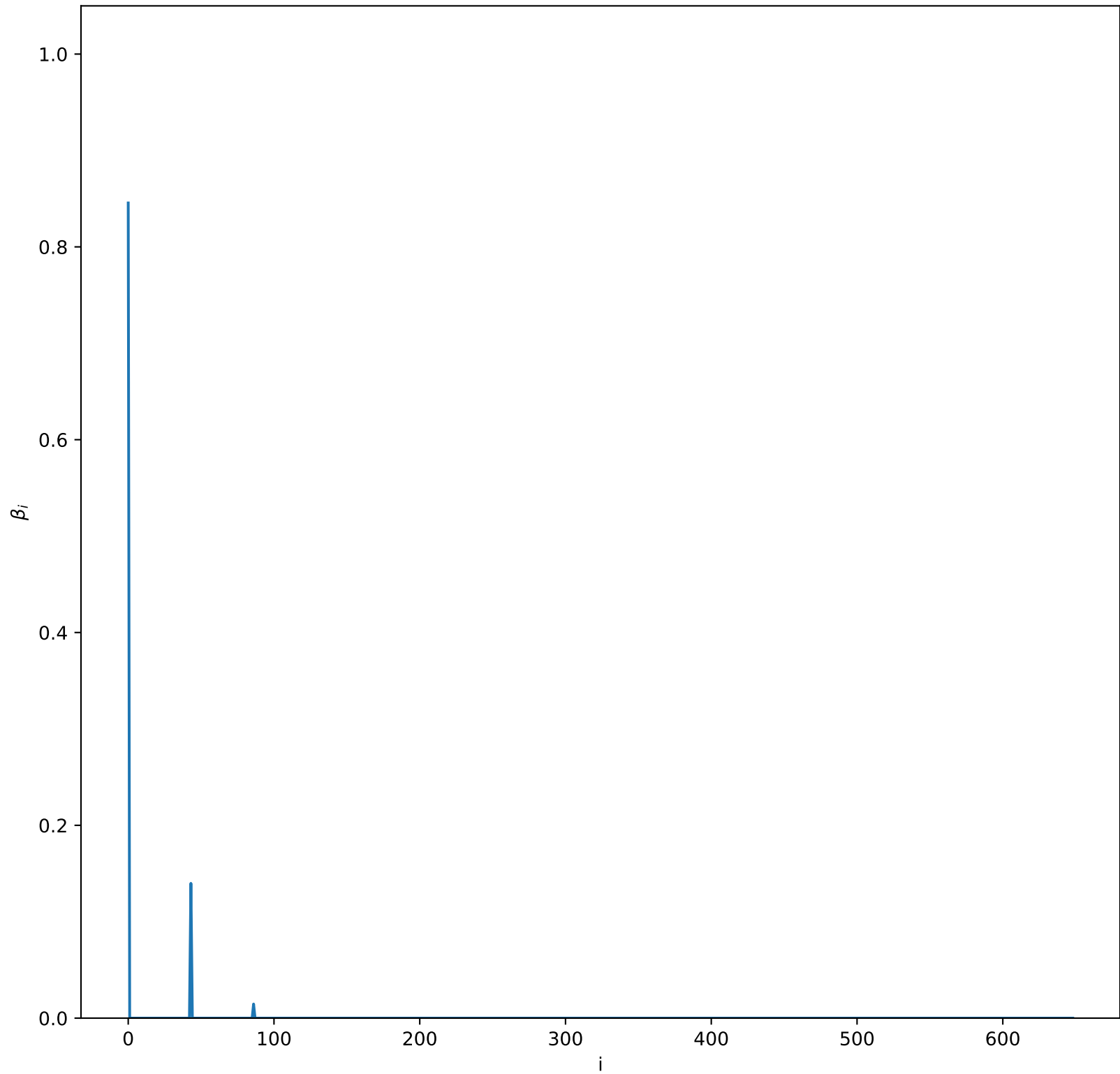
$\mu = 2.14$



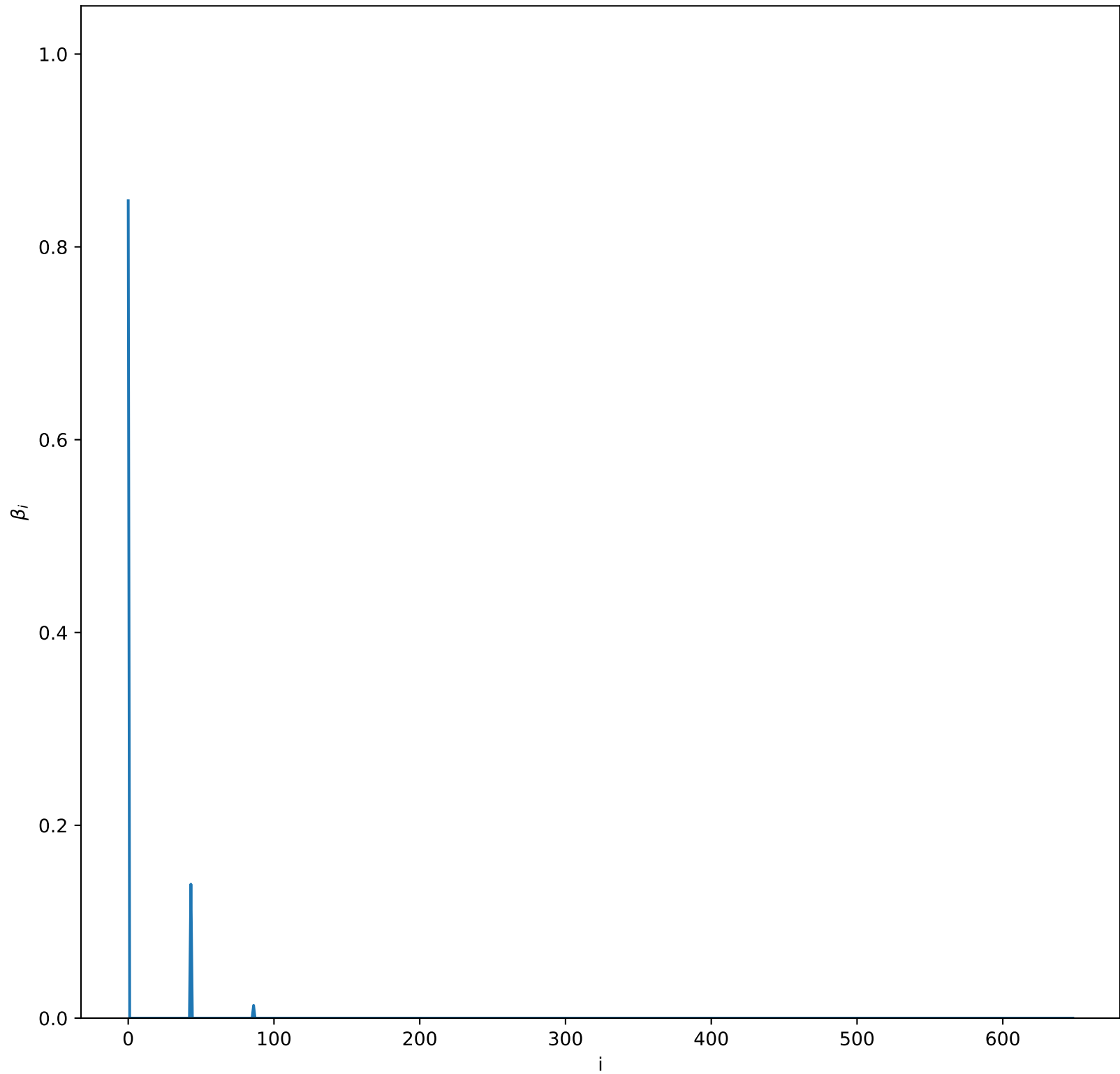
$\mu = 2.15$



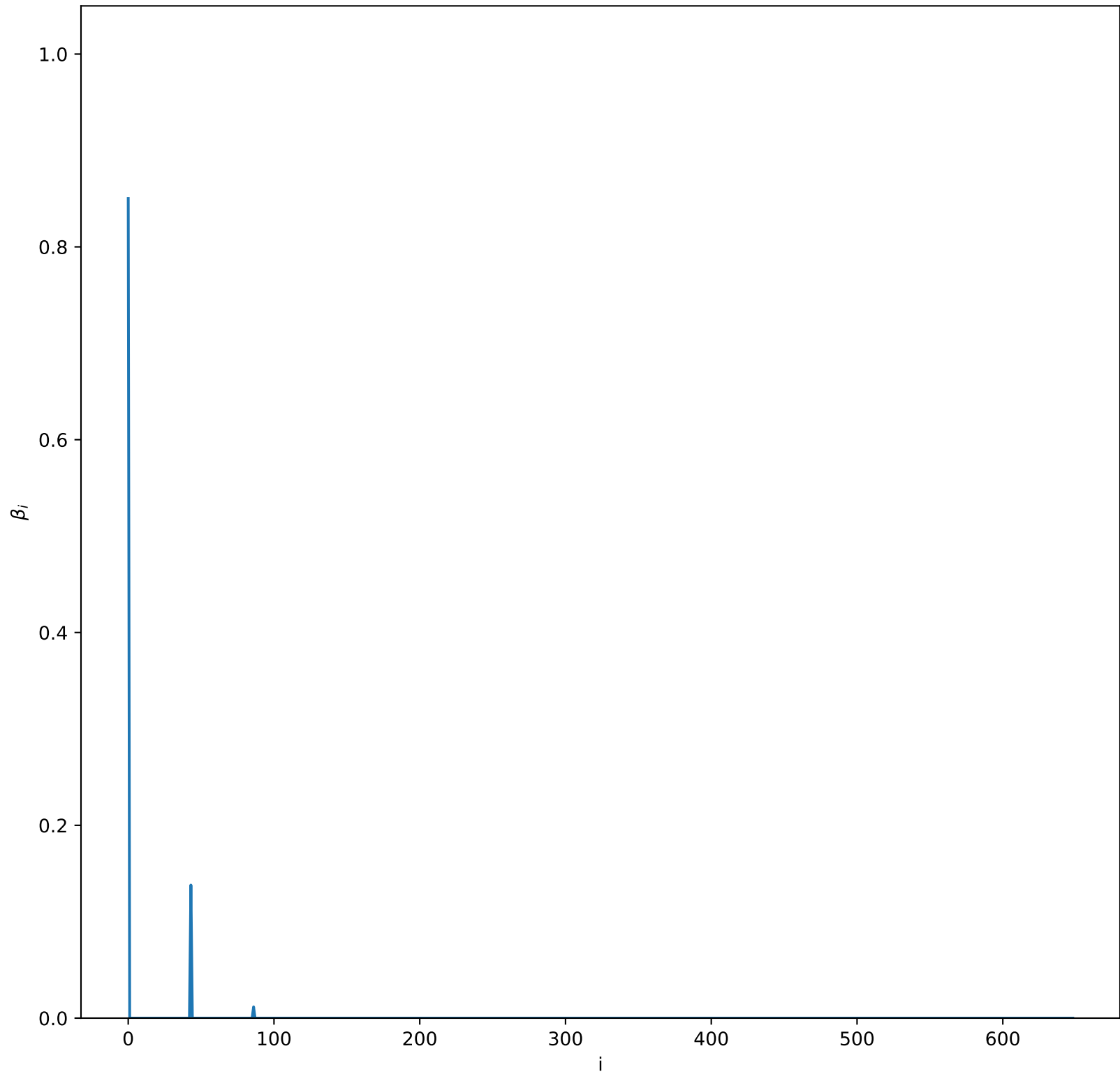
$\mu = 2.16$



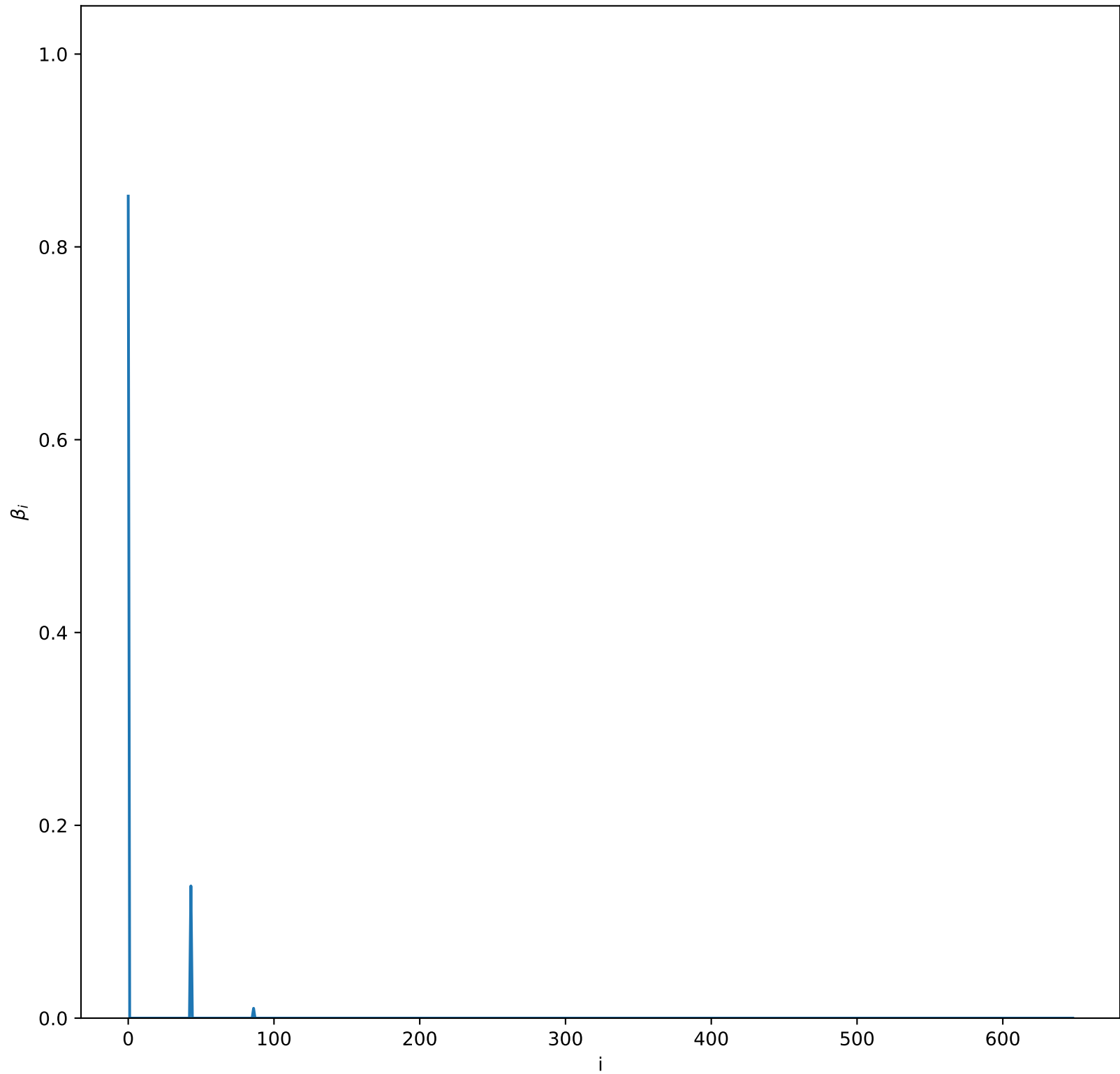
$\mu = 2.17$



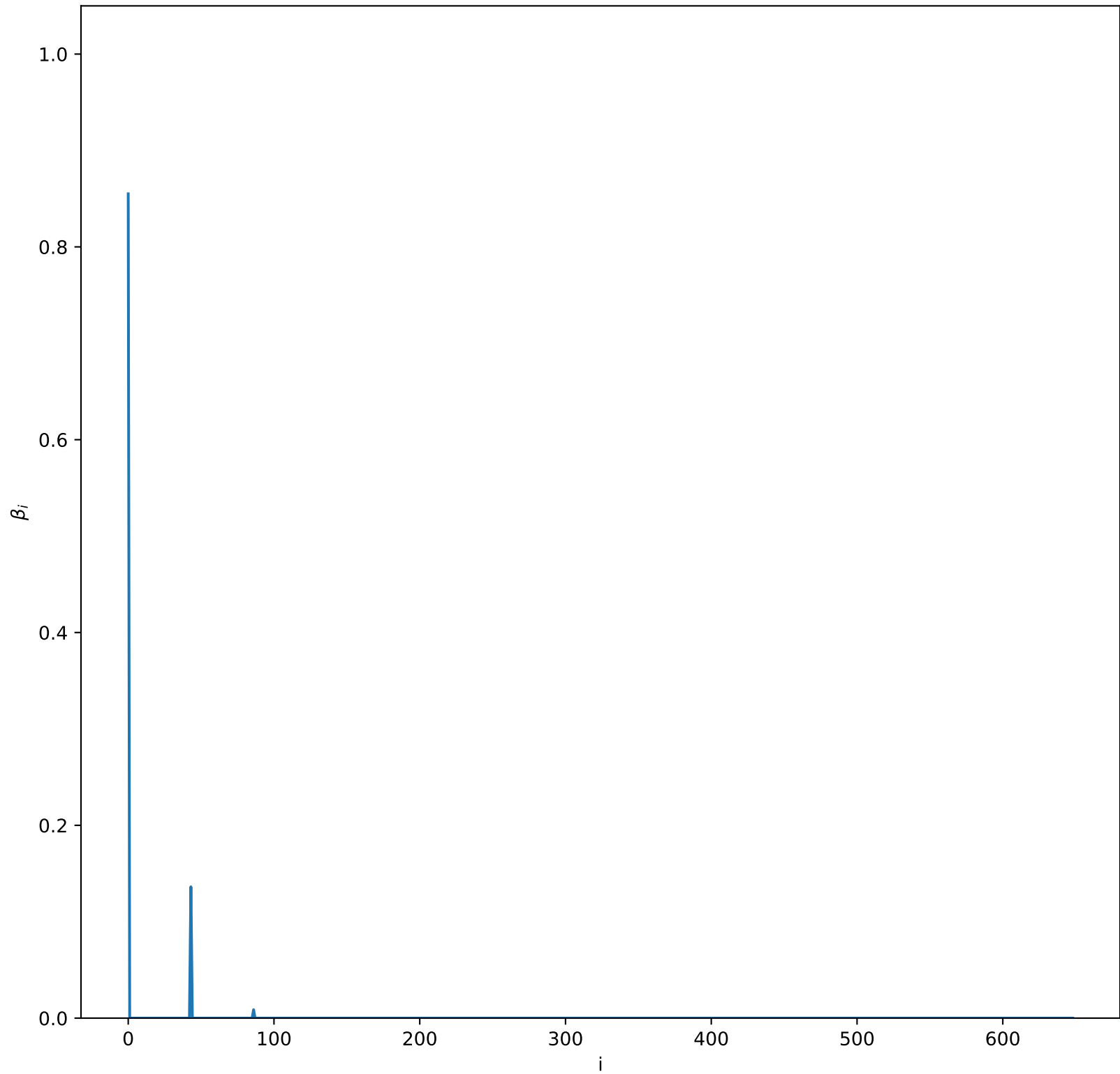
$\mu = 2.18$



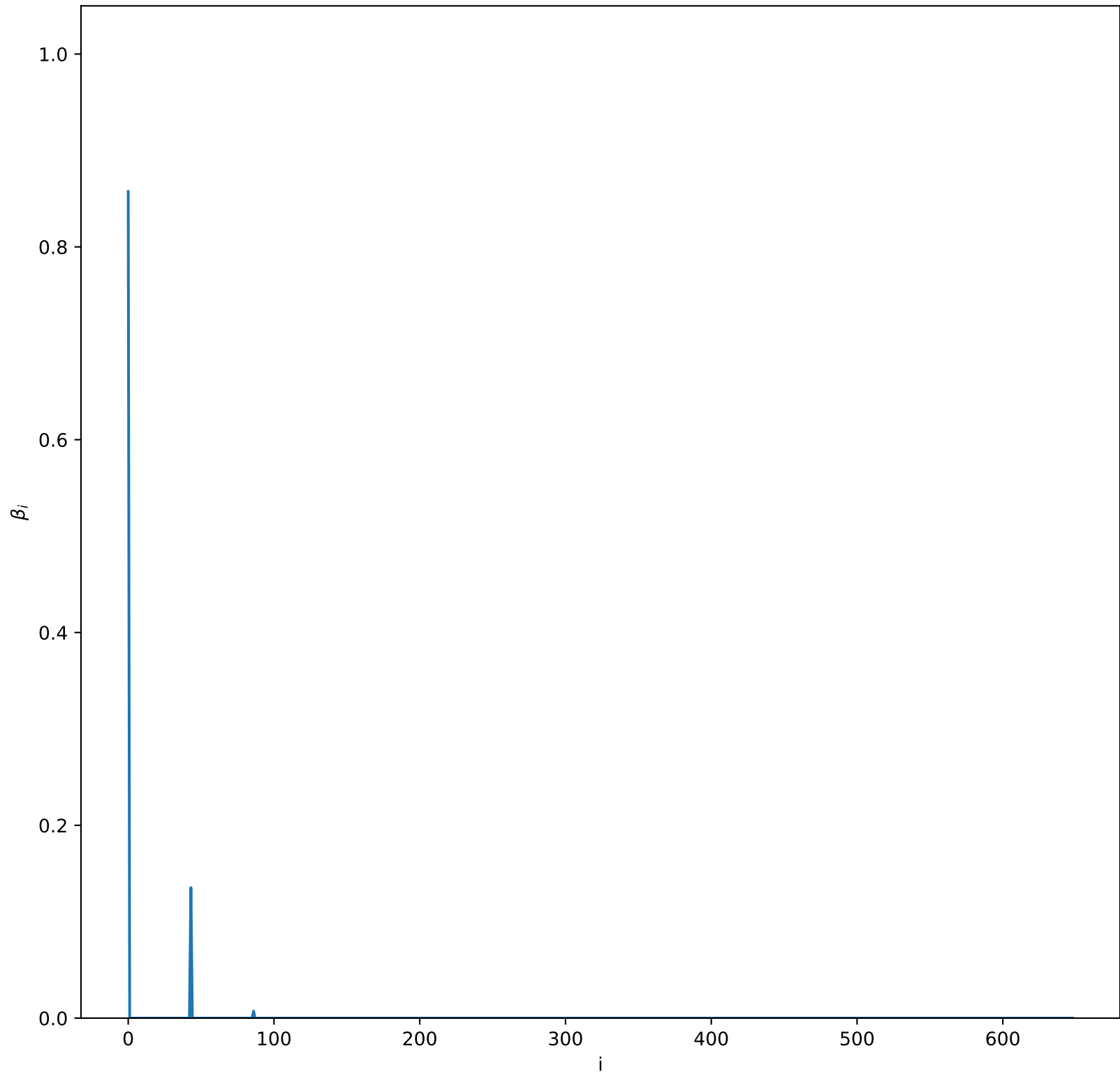
$\mu = 2.19$



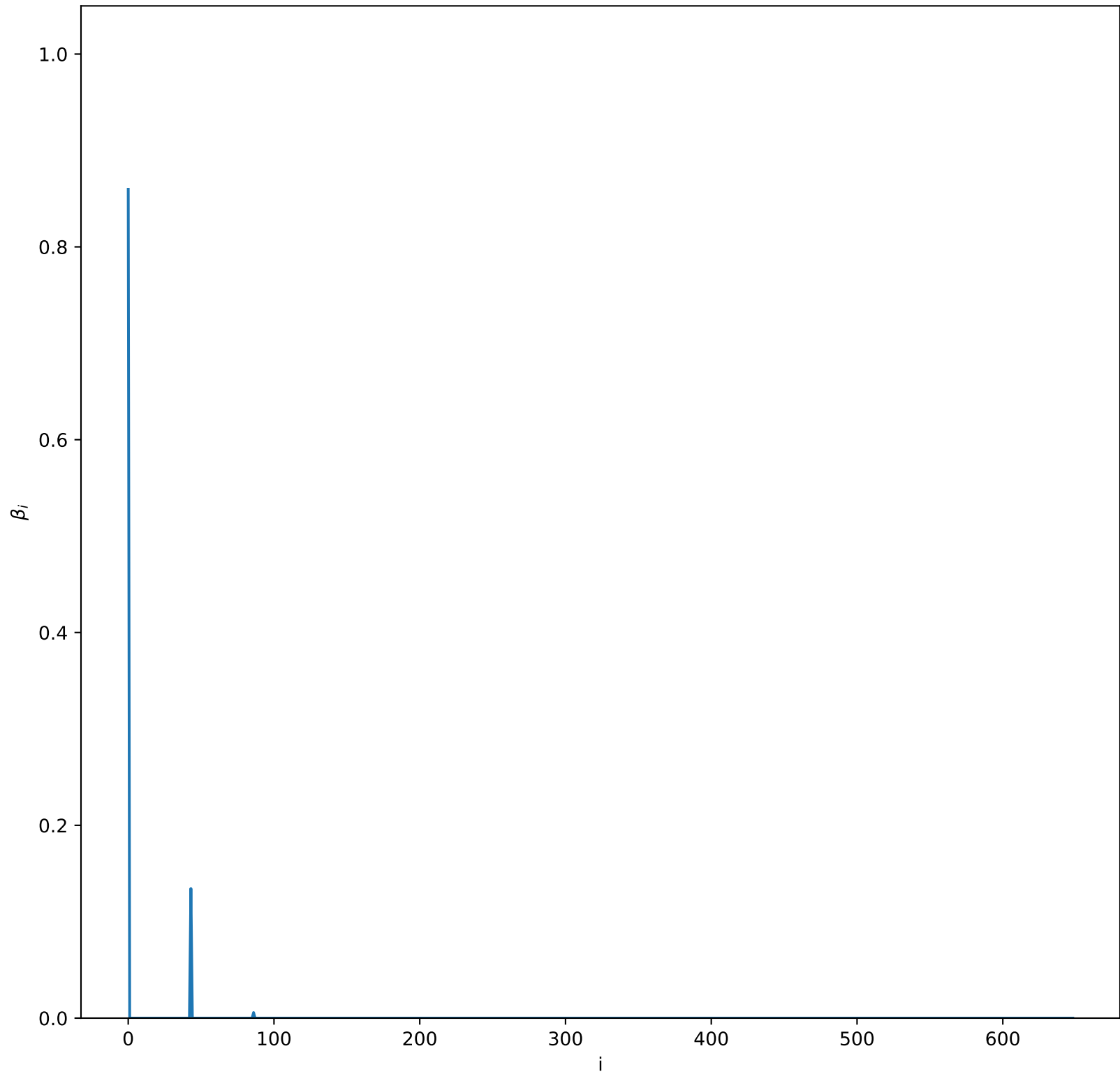
$\mu = 2.20$



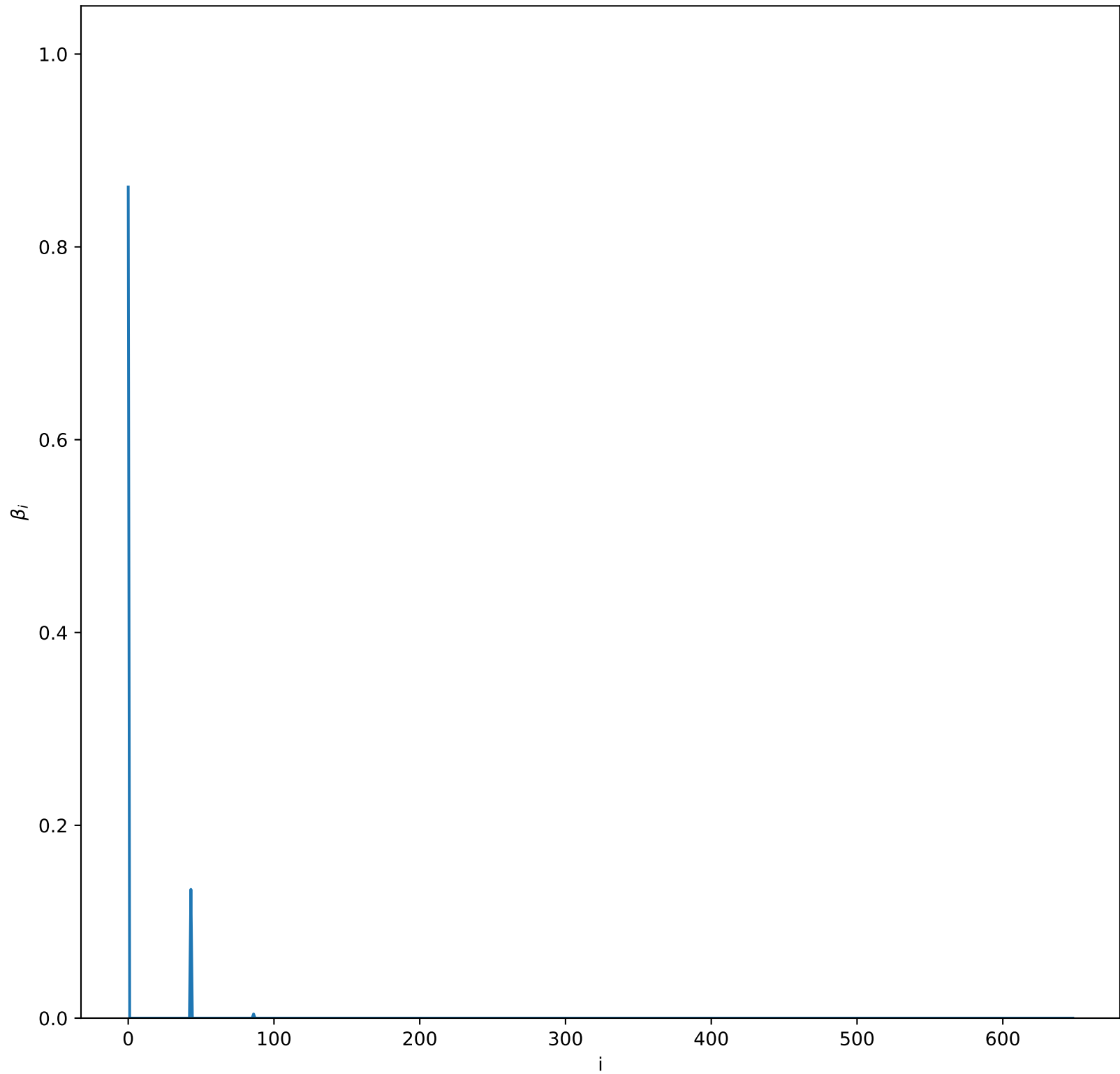
$\mu = 2.21$



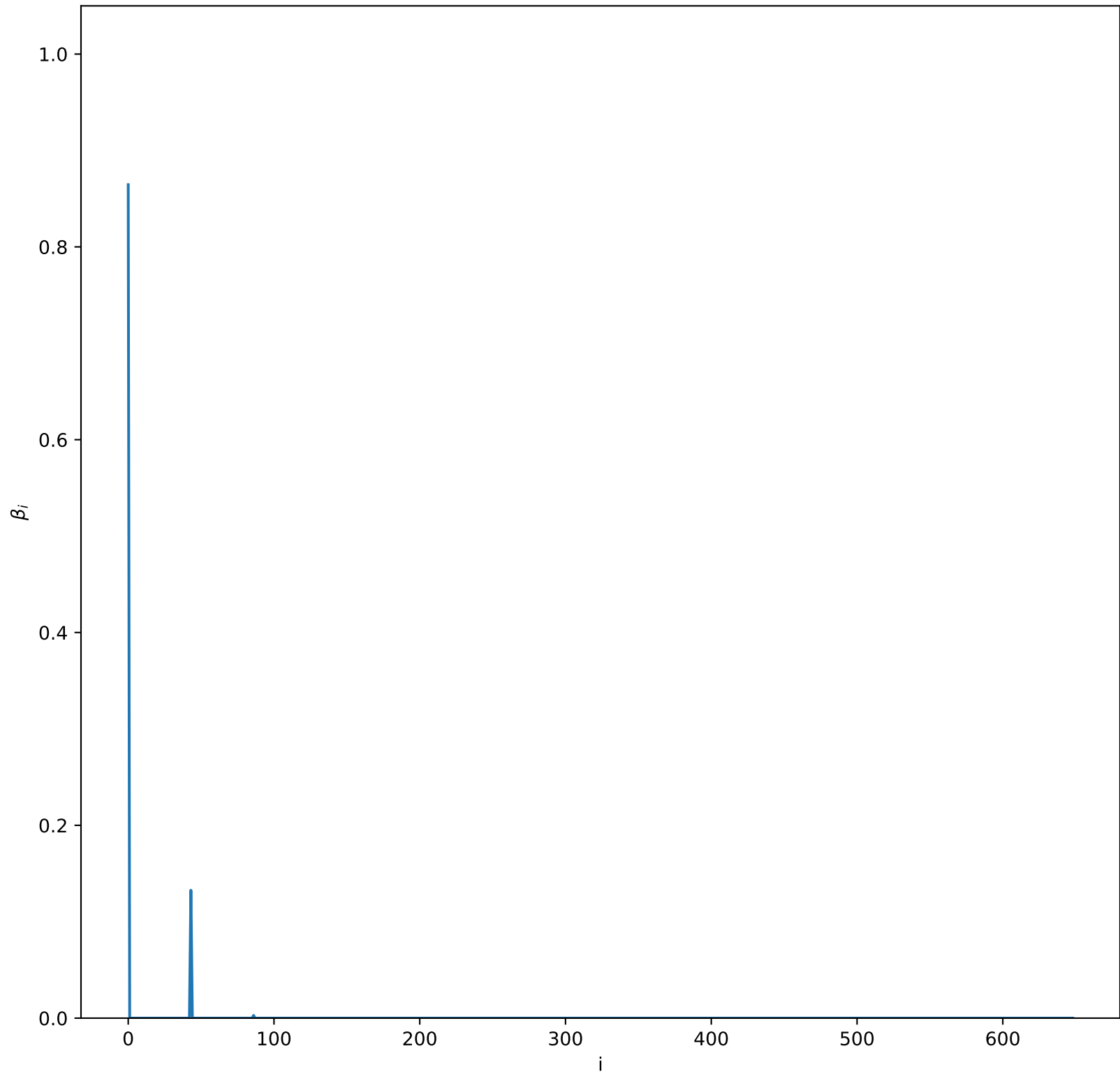
$\mu = 2.22$



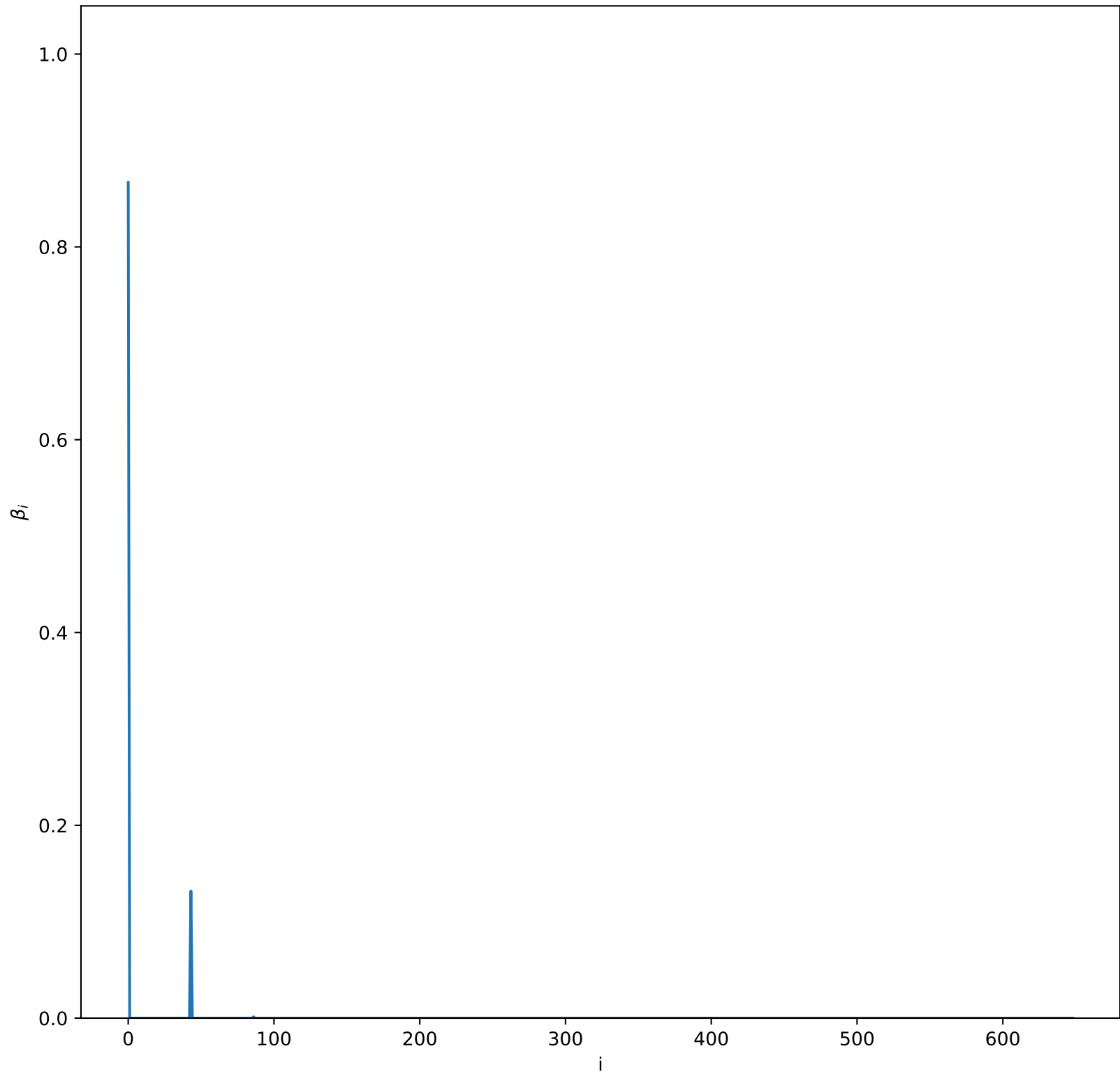
$\mu = 2.23$



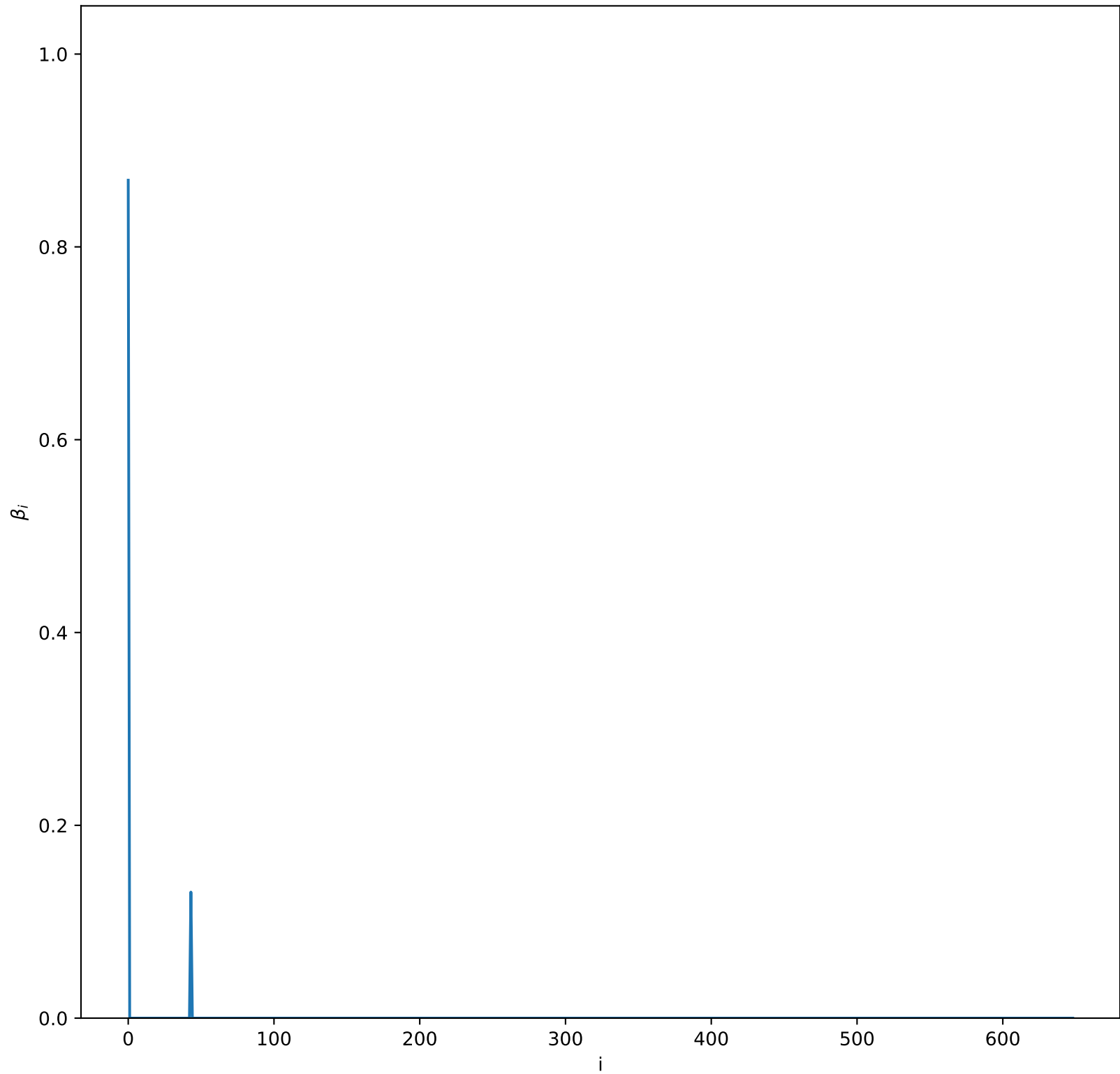
$\mu = 2.24$



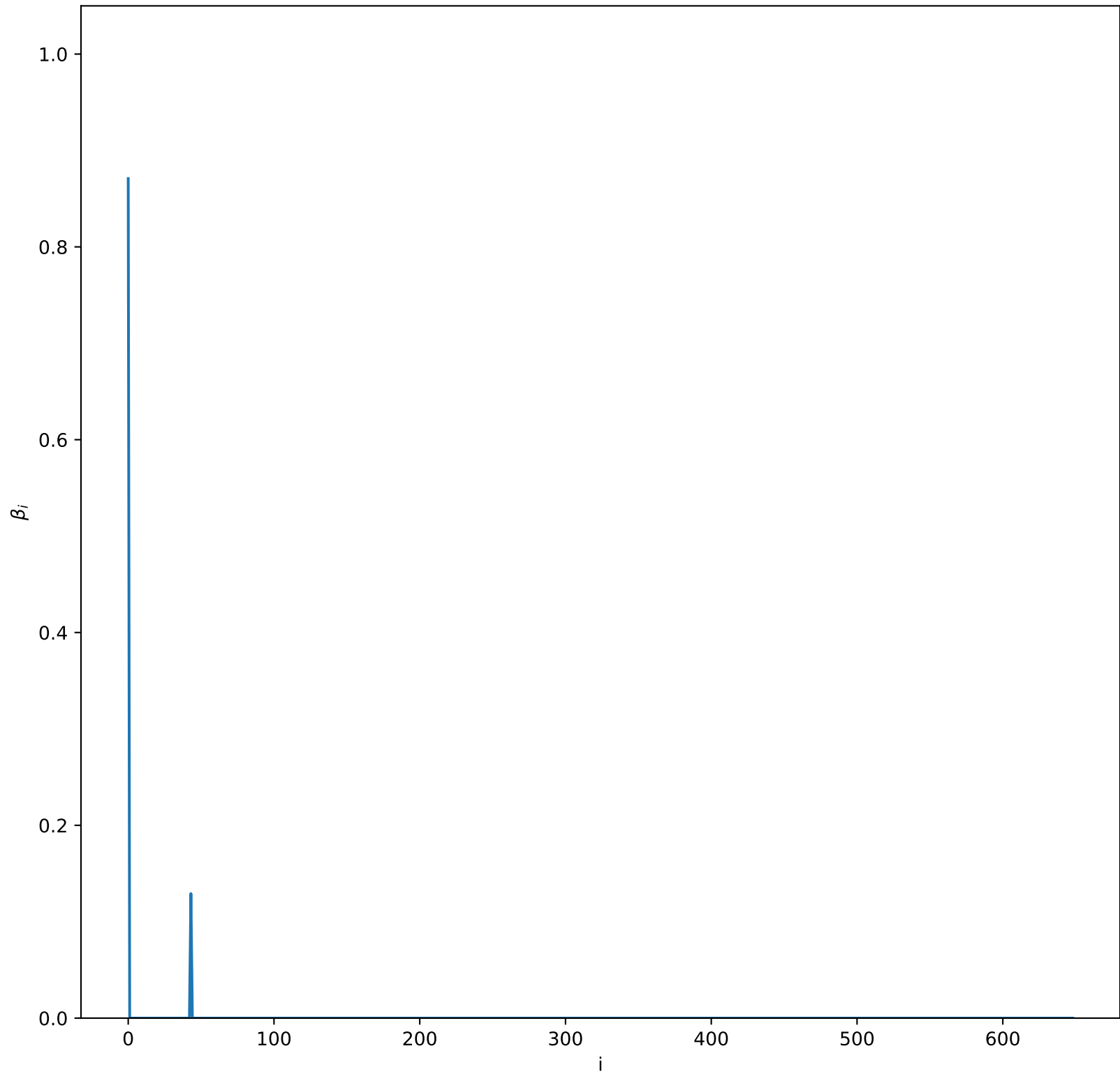
$\mu = 2.25$



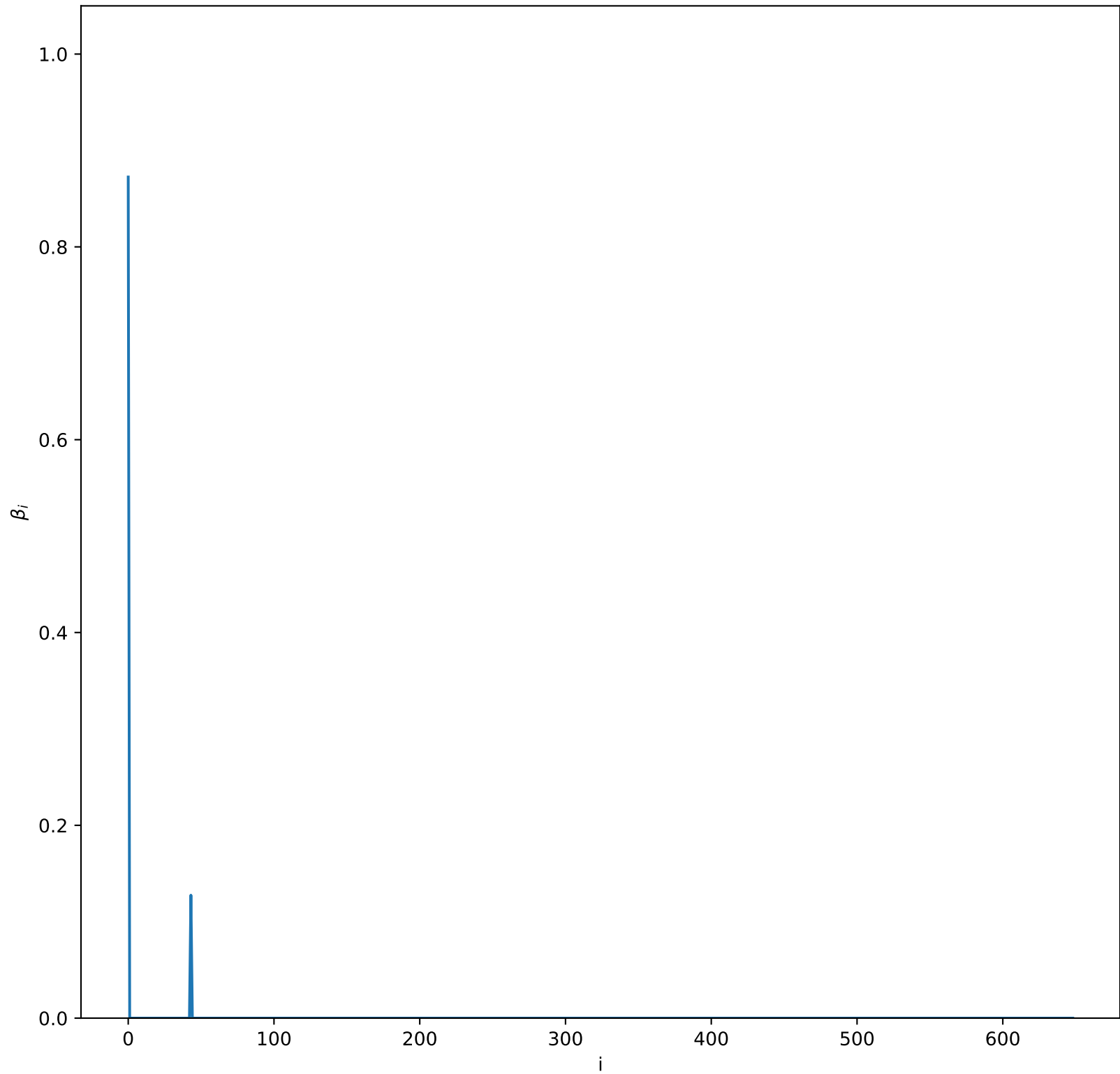
$\mu = 2.26$



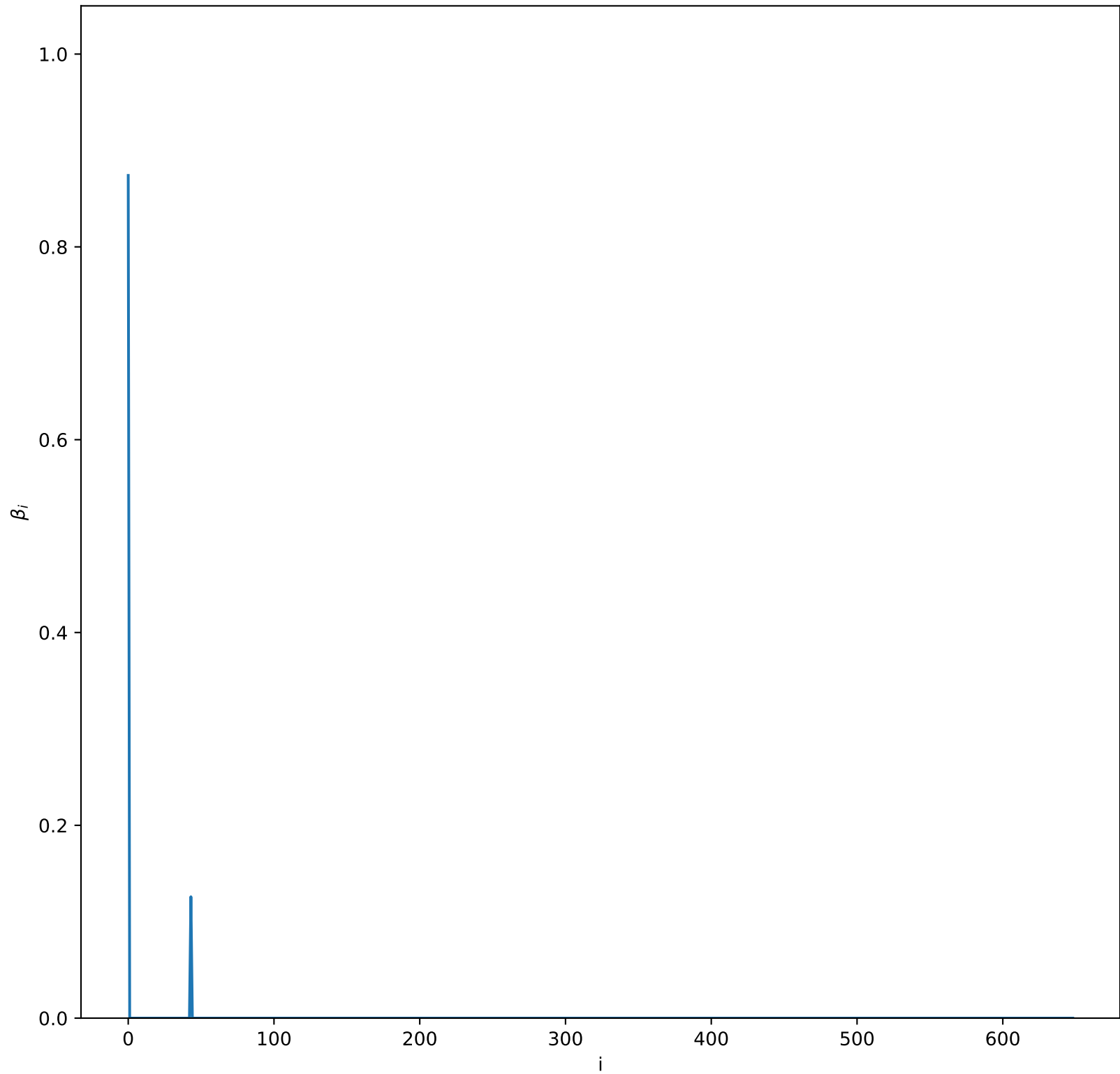
$\mu = 2.27$



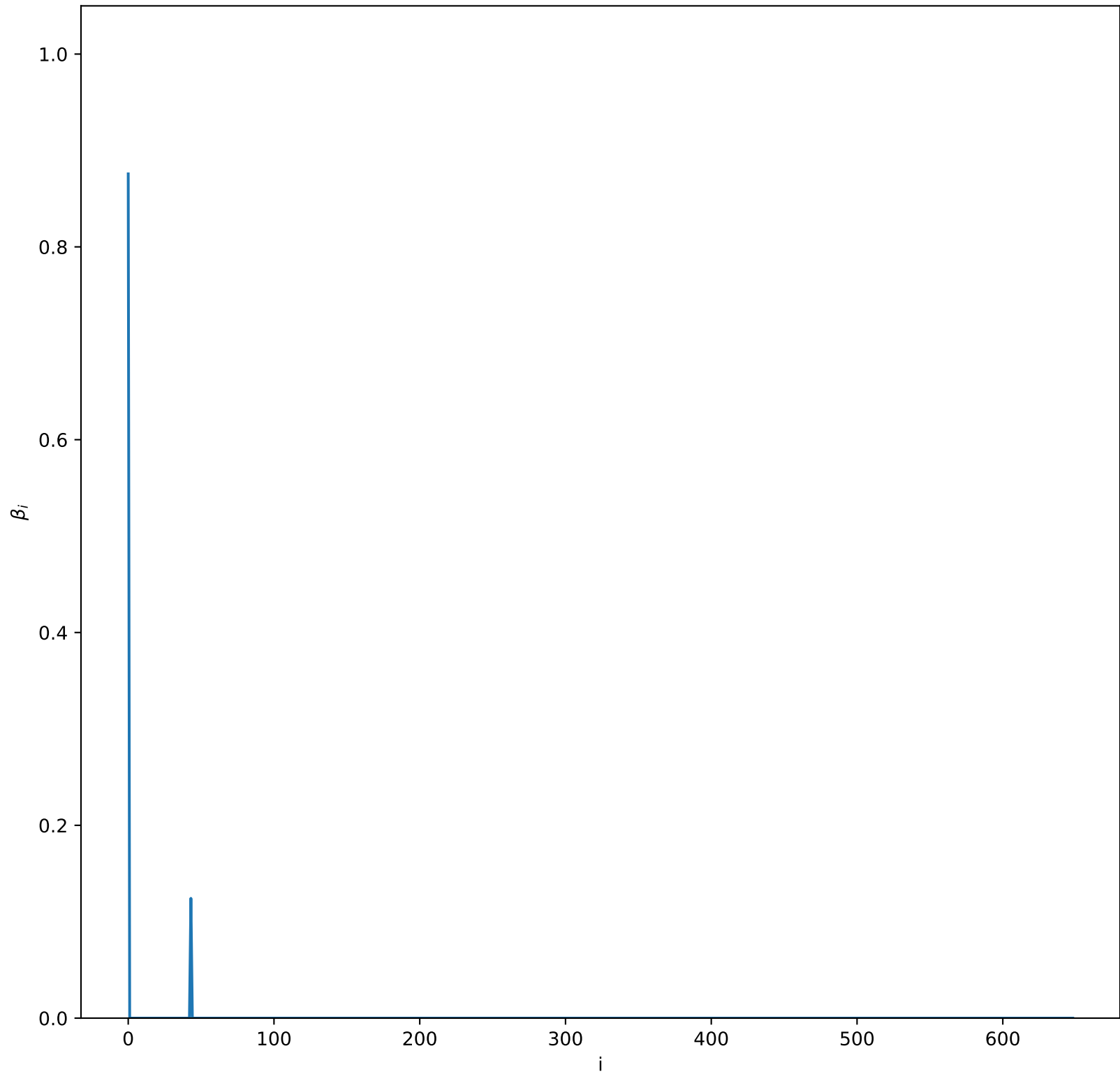
$\mu = 2.28$



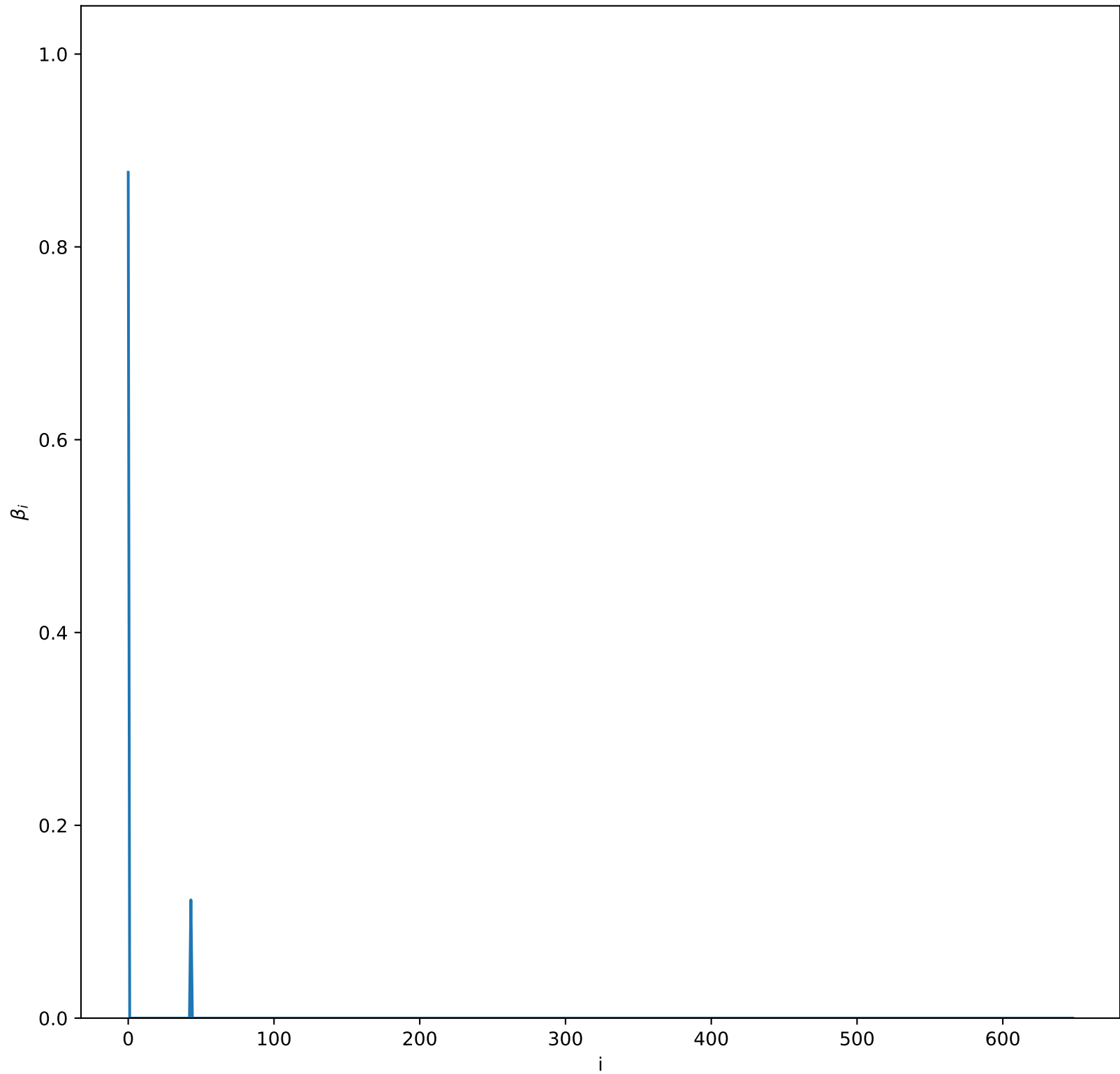
$\mu = 2.29$



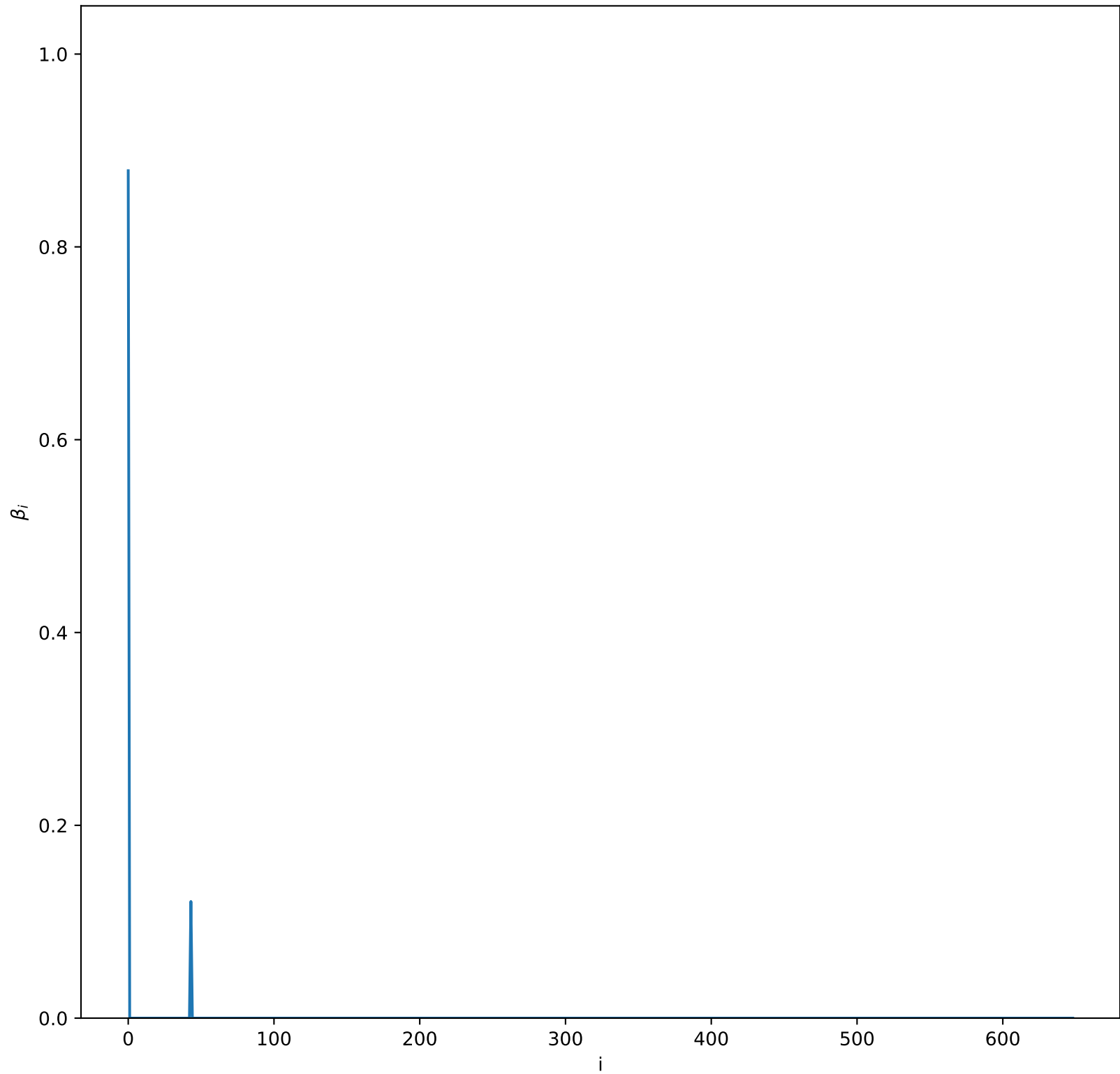
$\mu = 2.30$



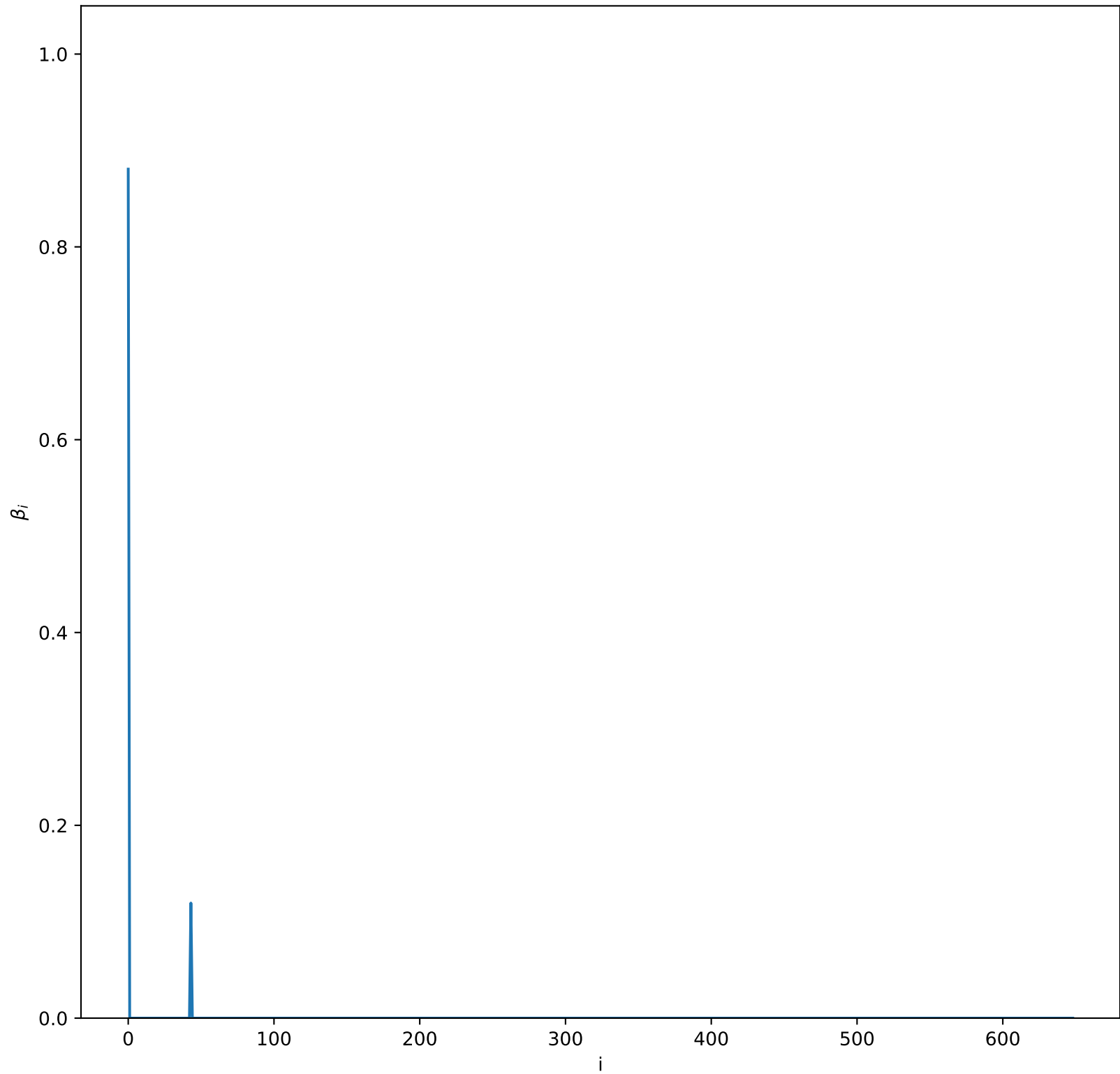
$\mu = 2.31$



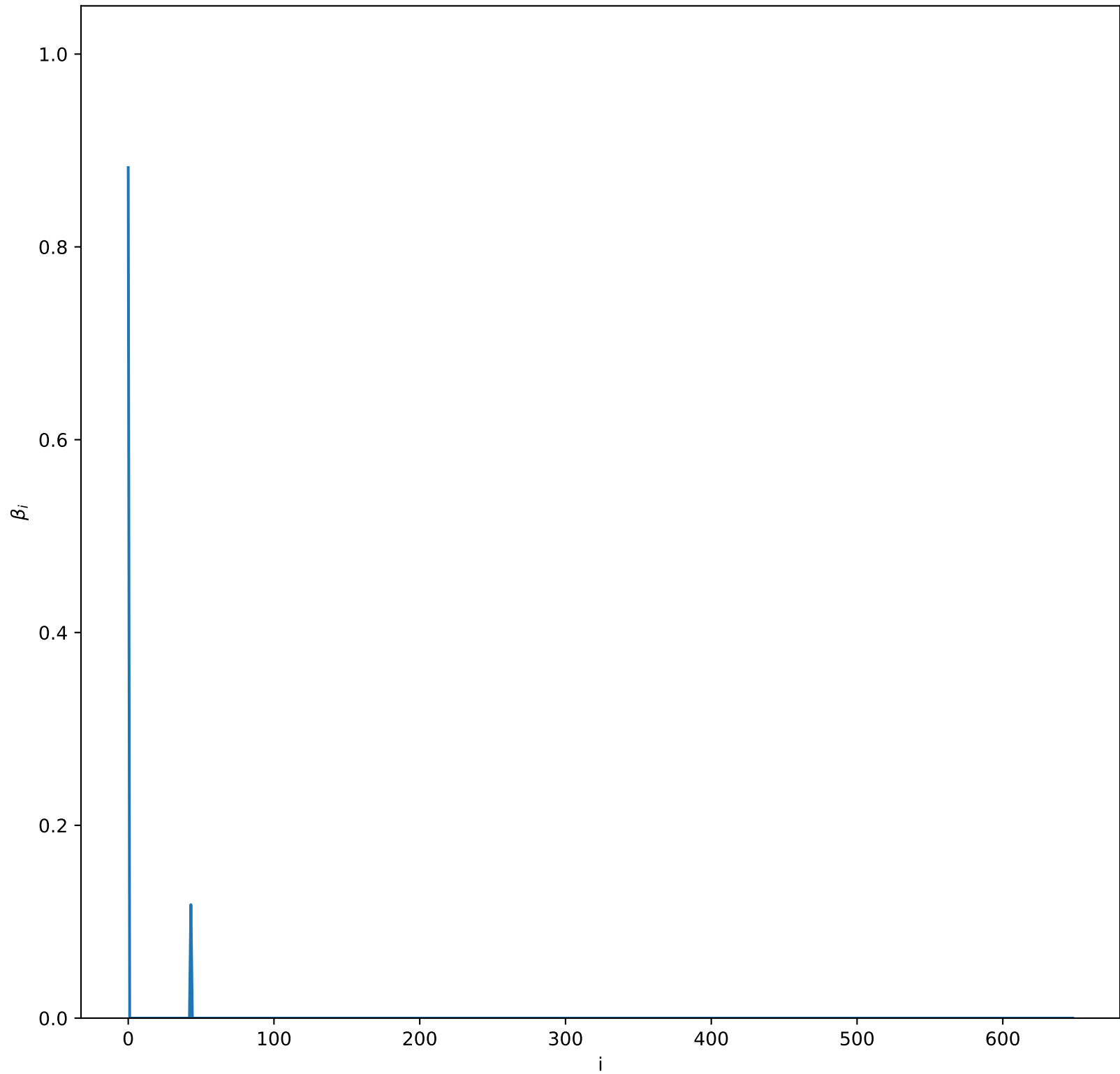
$\mu = 2.32$



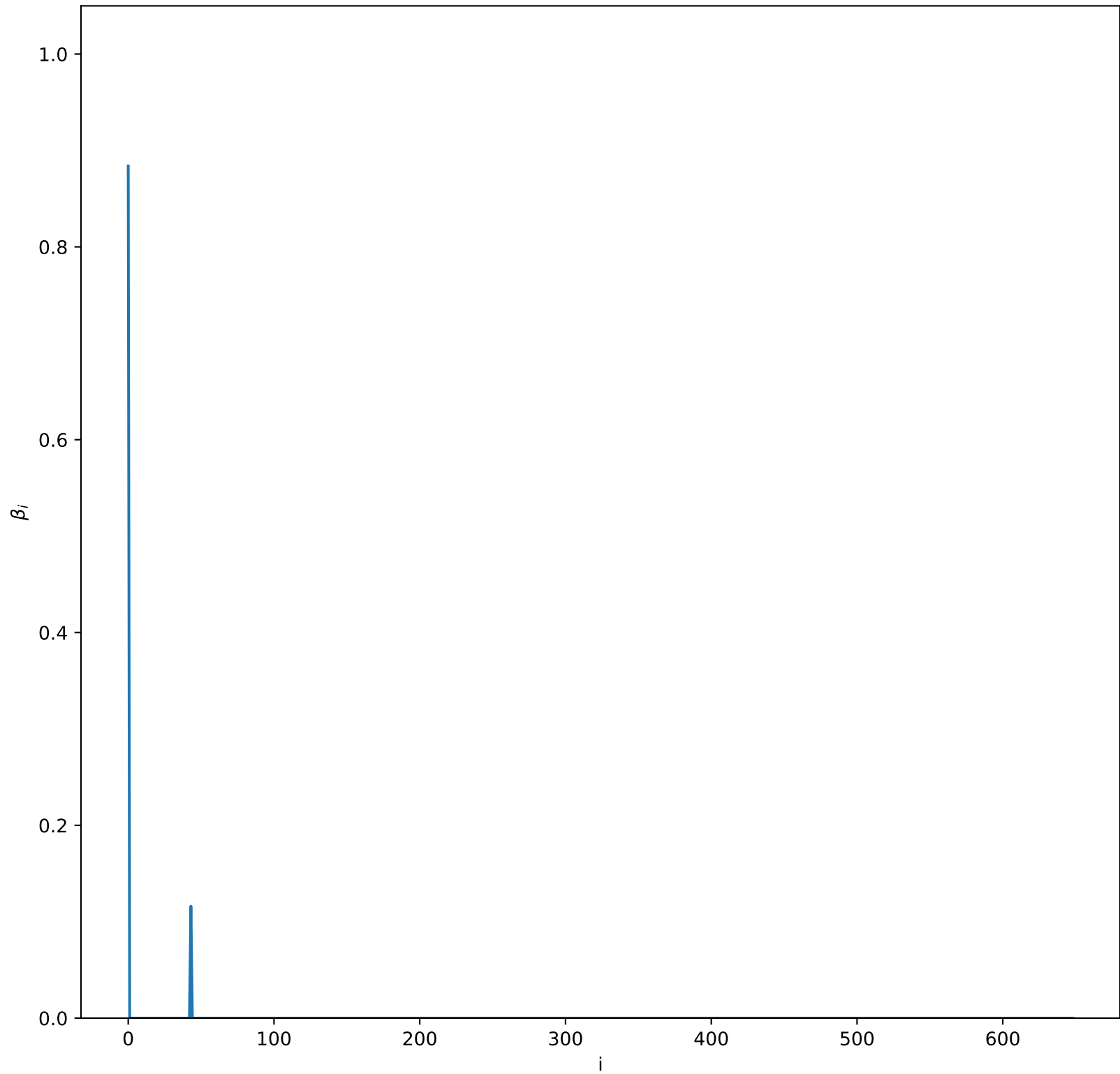
$\mu = 2.33$



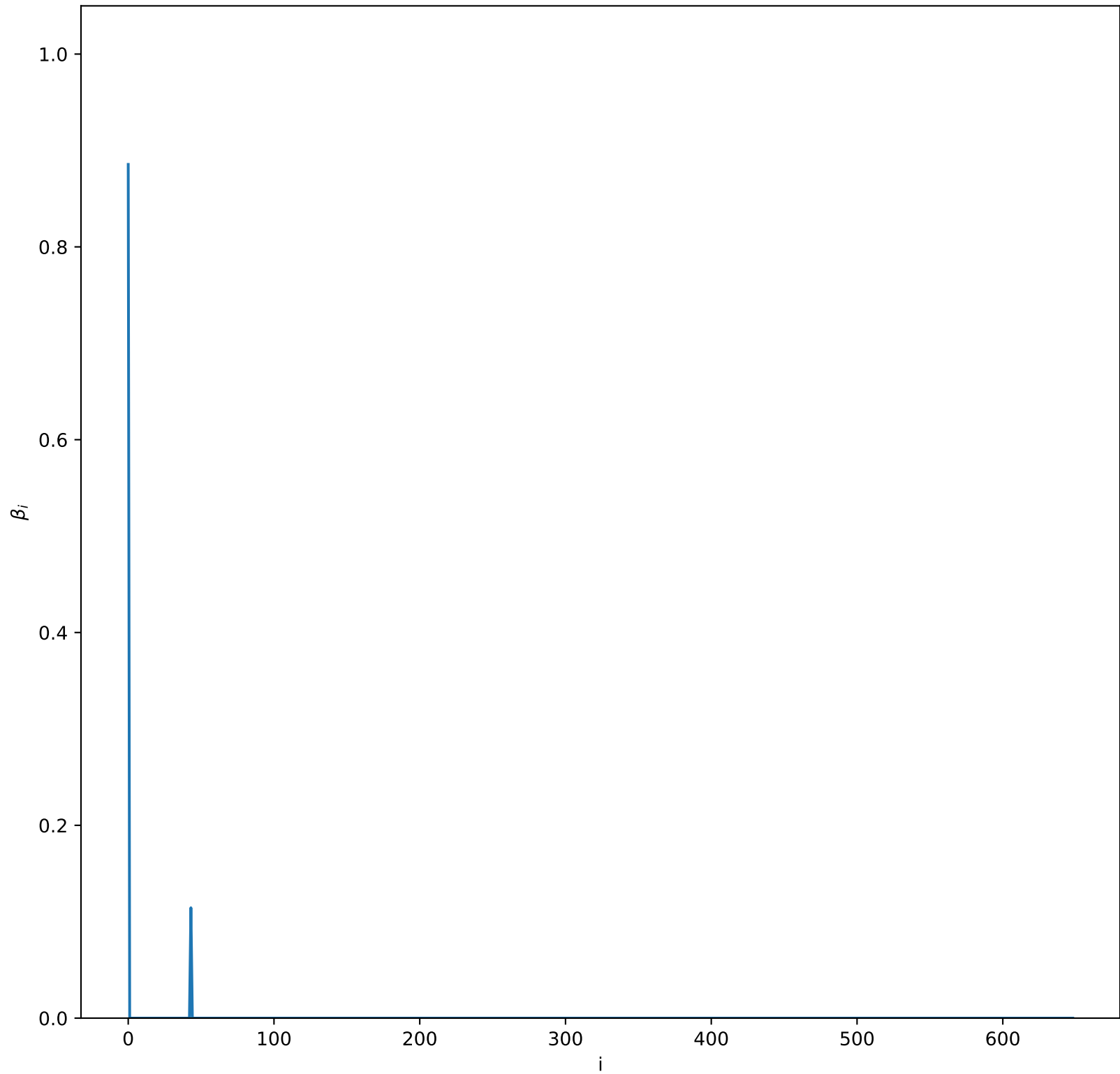
$\mu = 2.34$



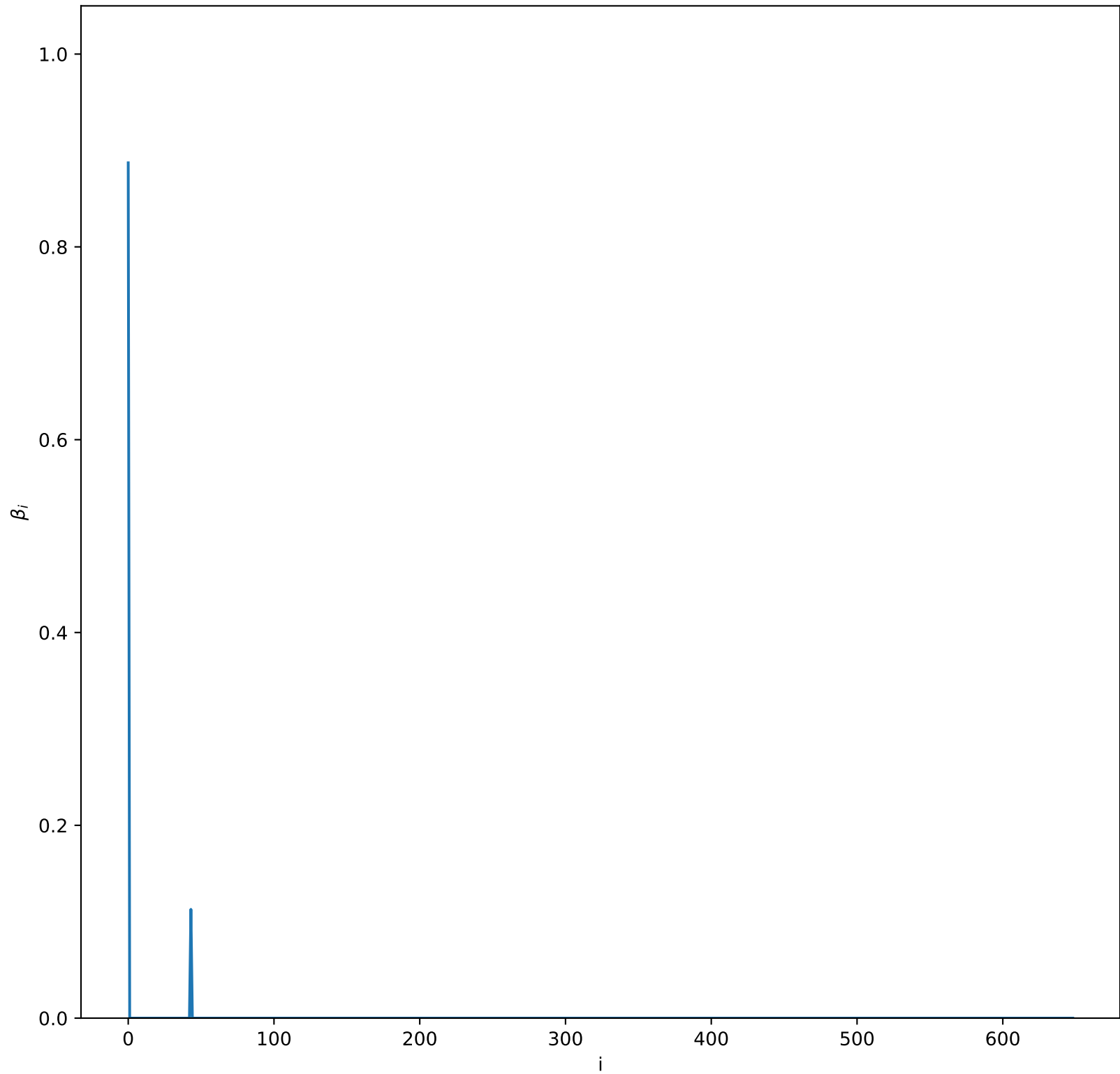
$\mu = 2.35$



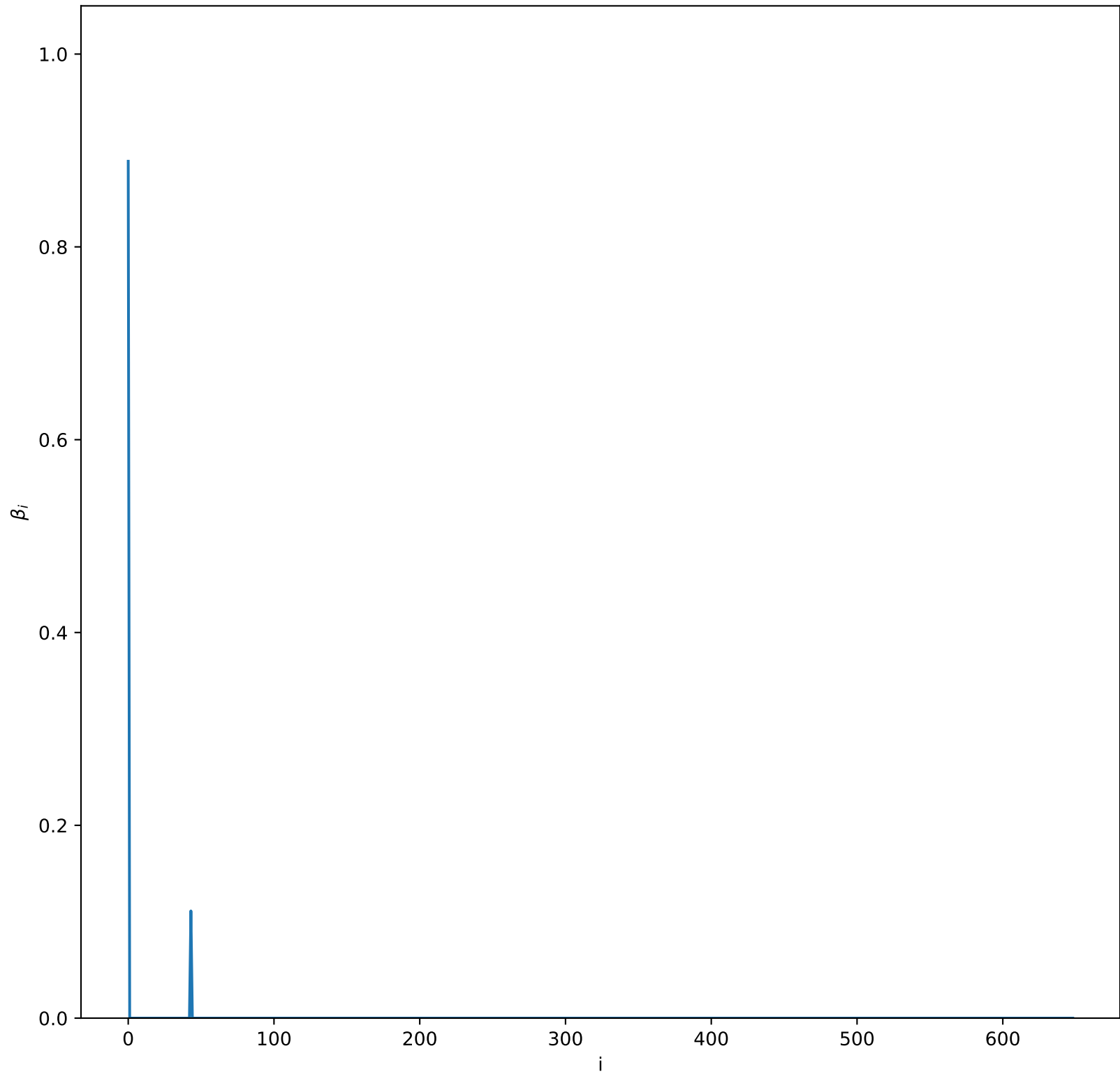
$\mu = 2.36$



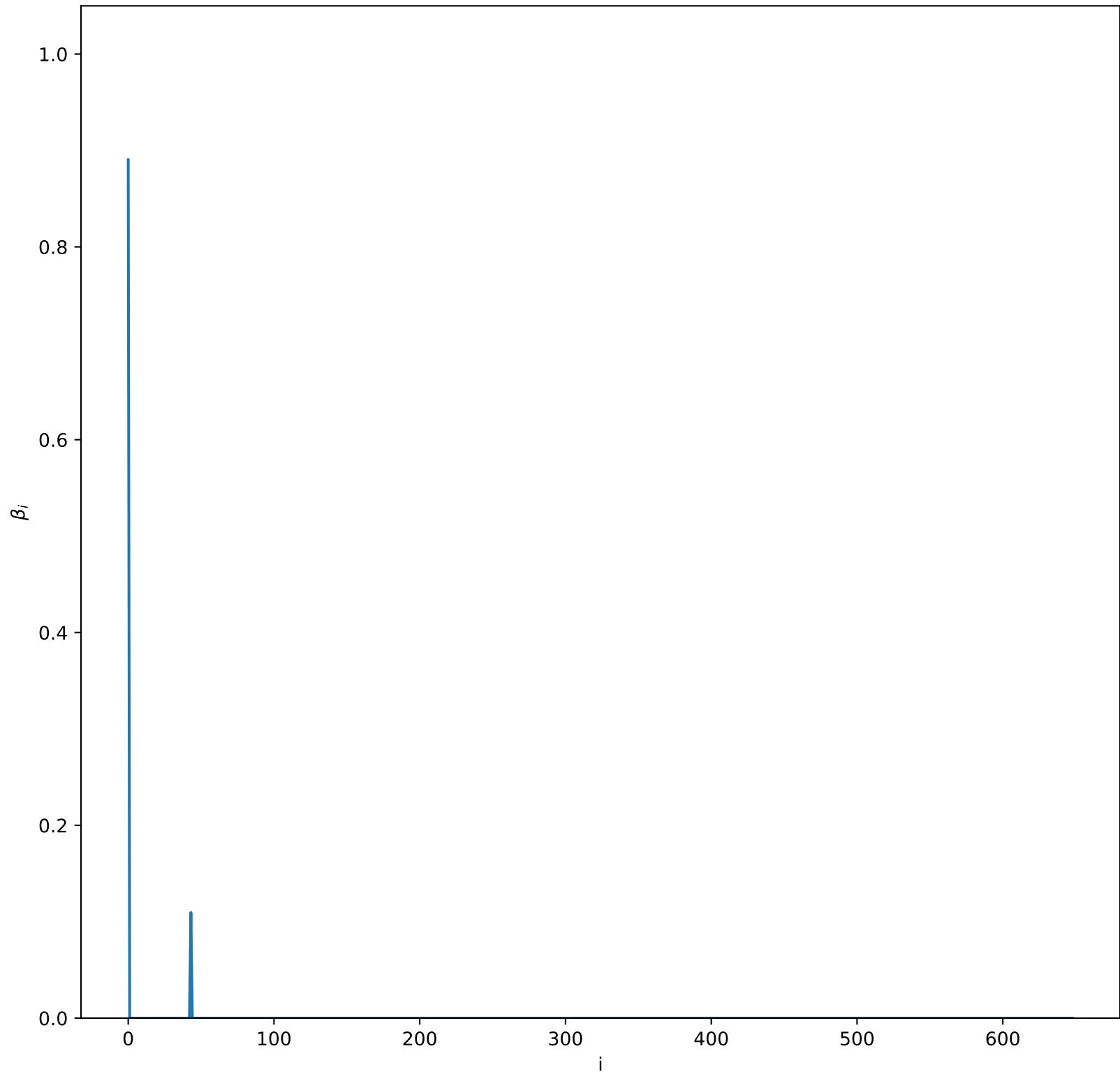
$\mu = 2.37$



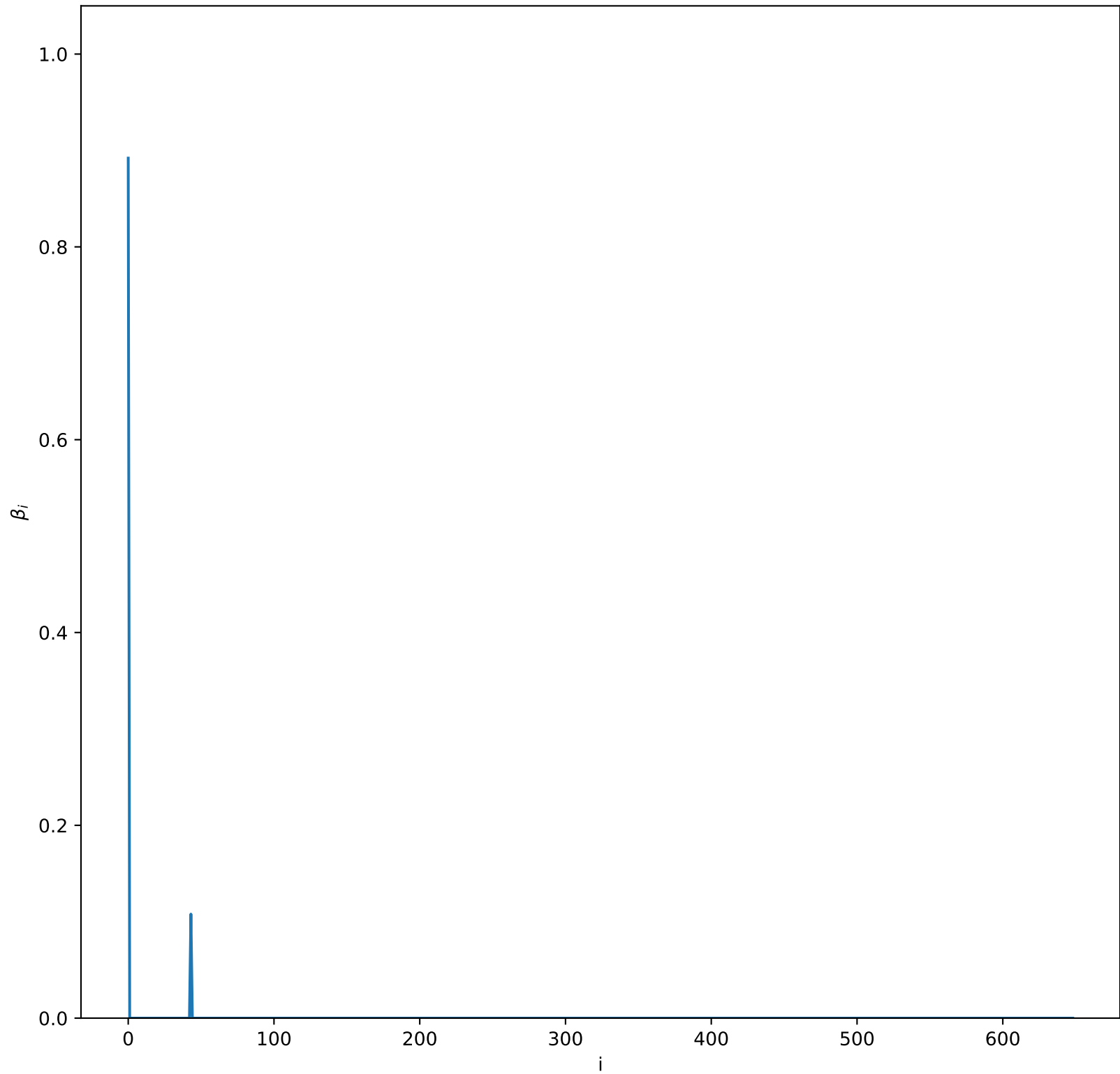
$\mu = 2.38$



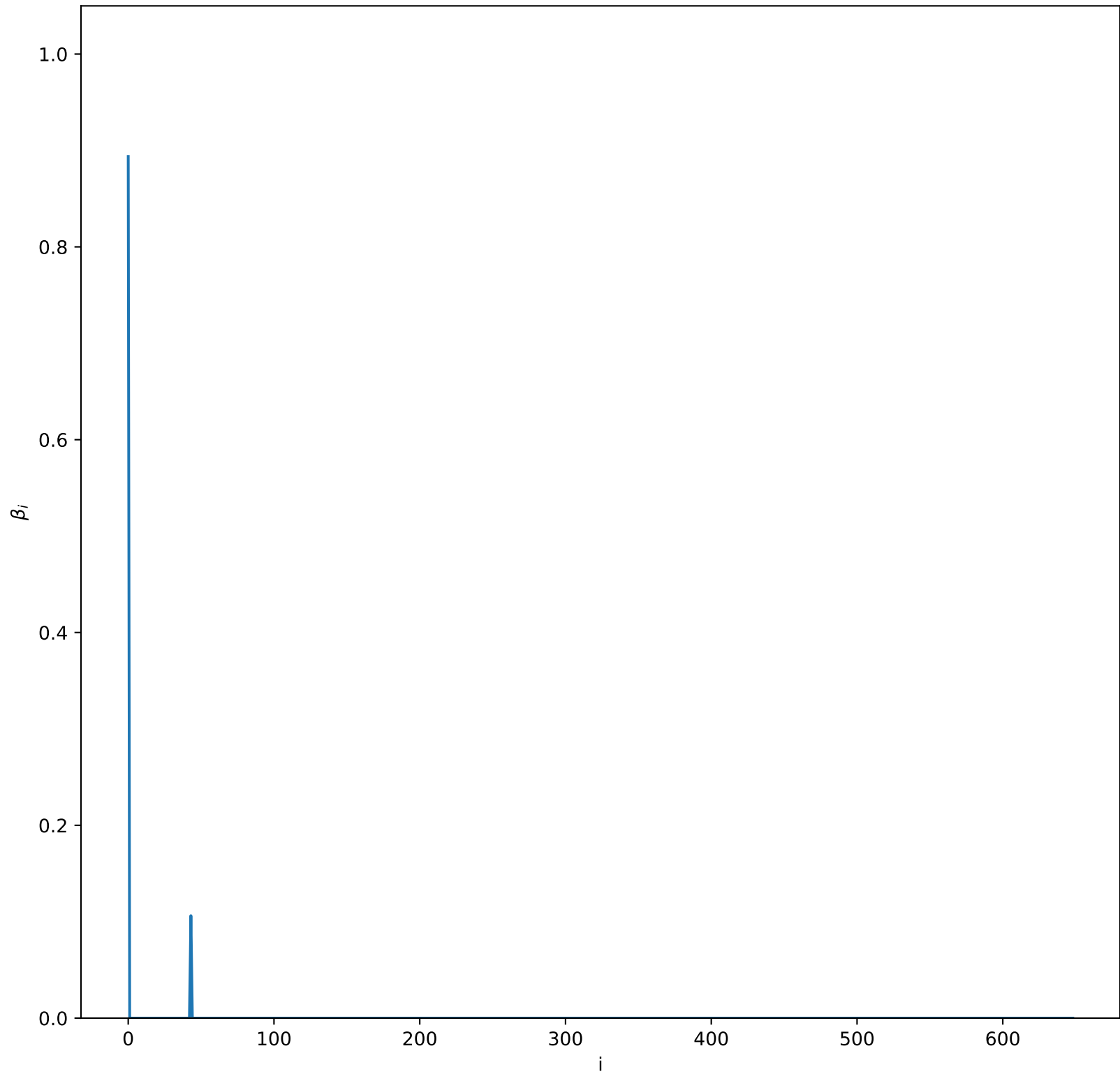
$\mu = 2.39$



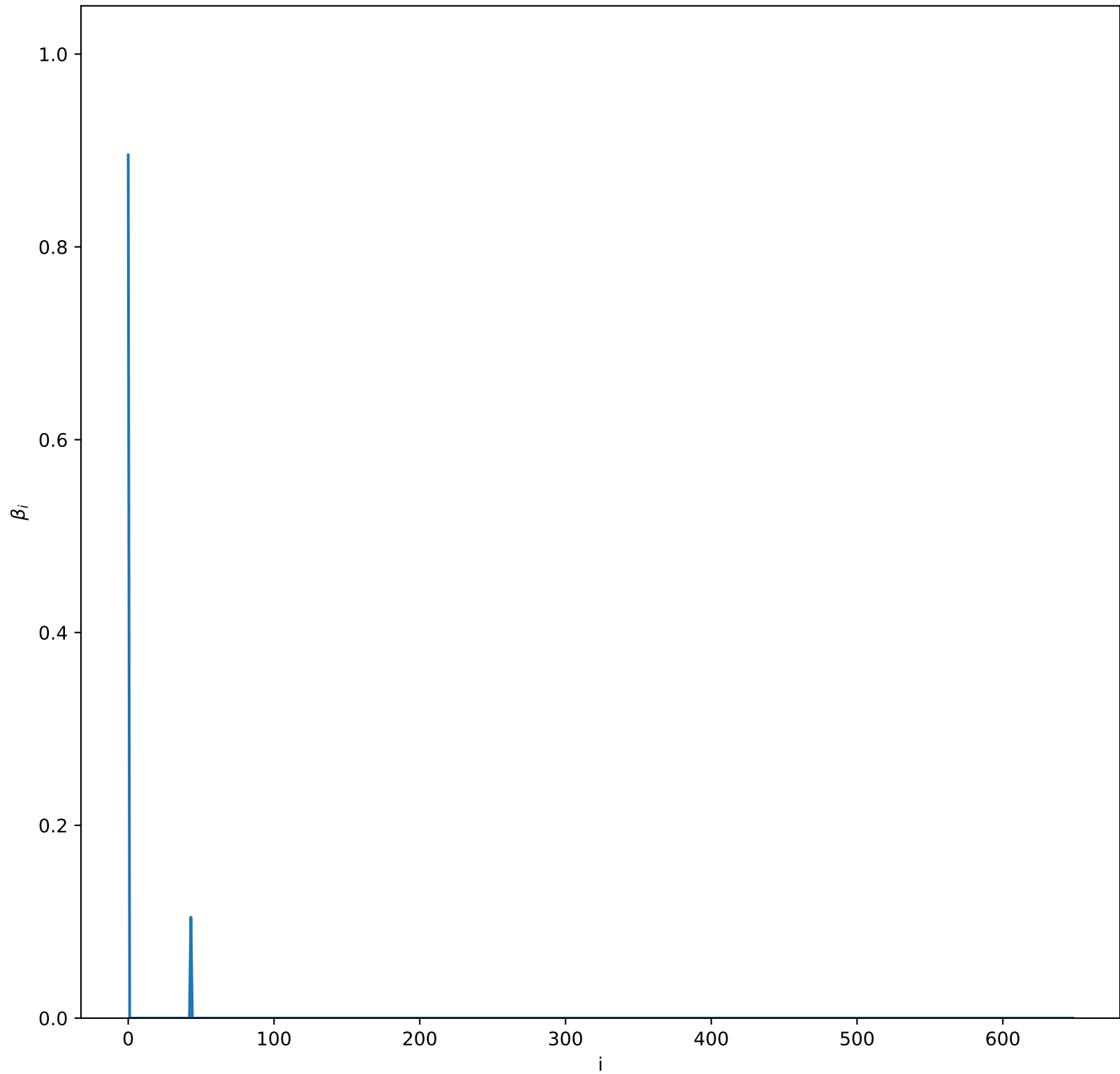
$\mu = 2.40$



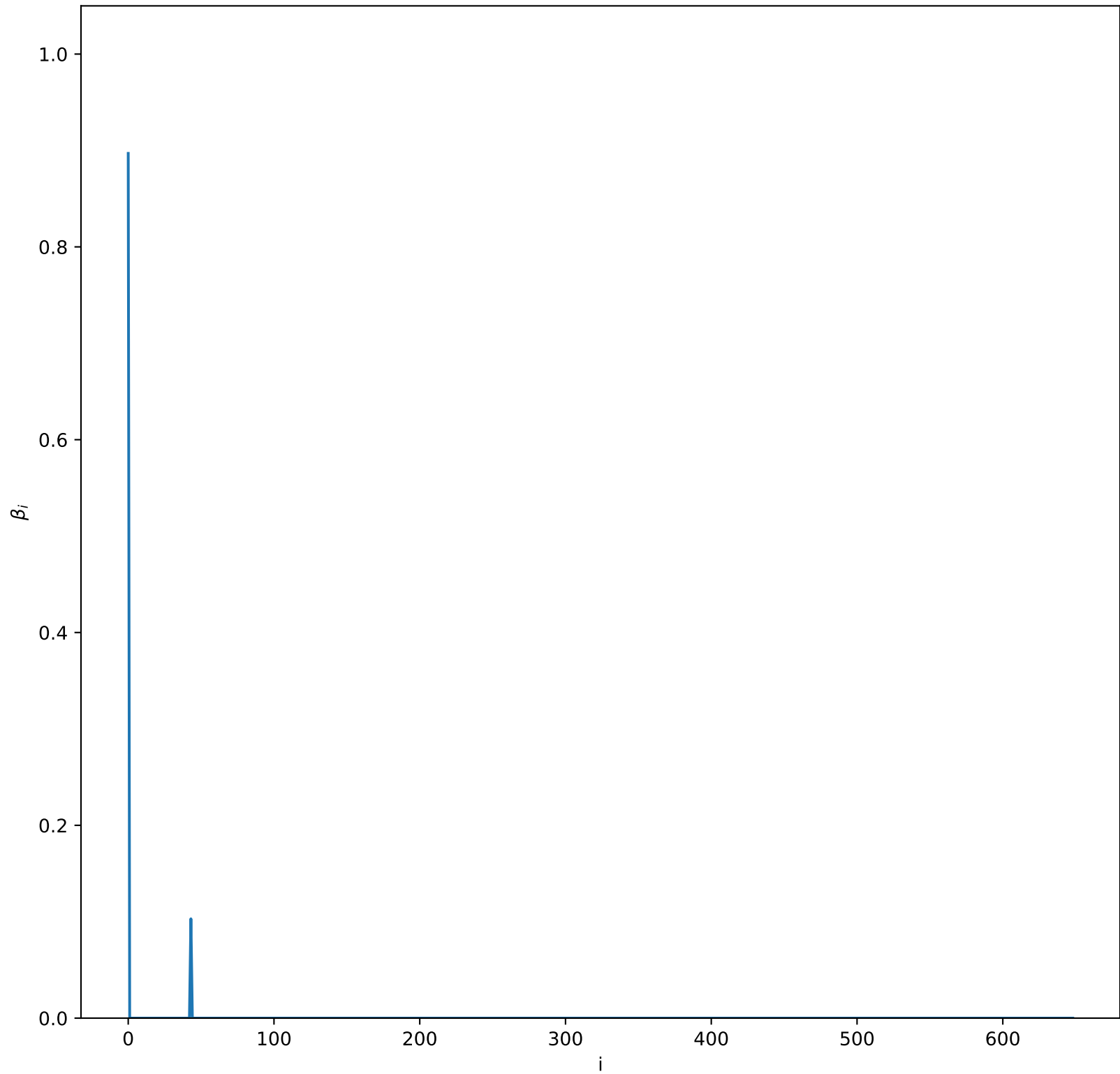
$\mu = 2.41$



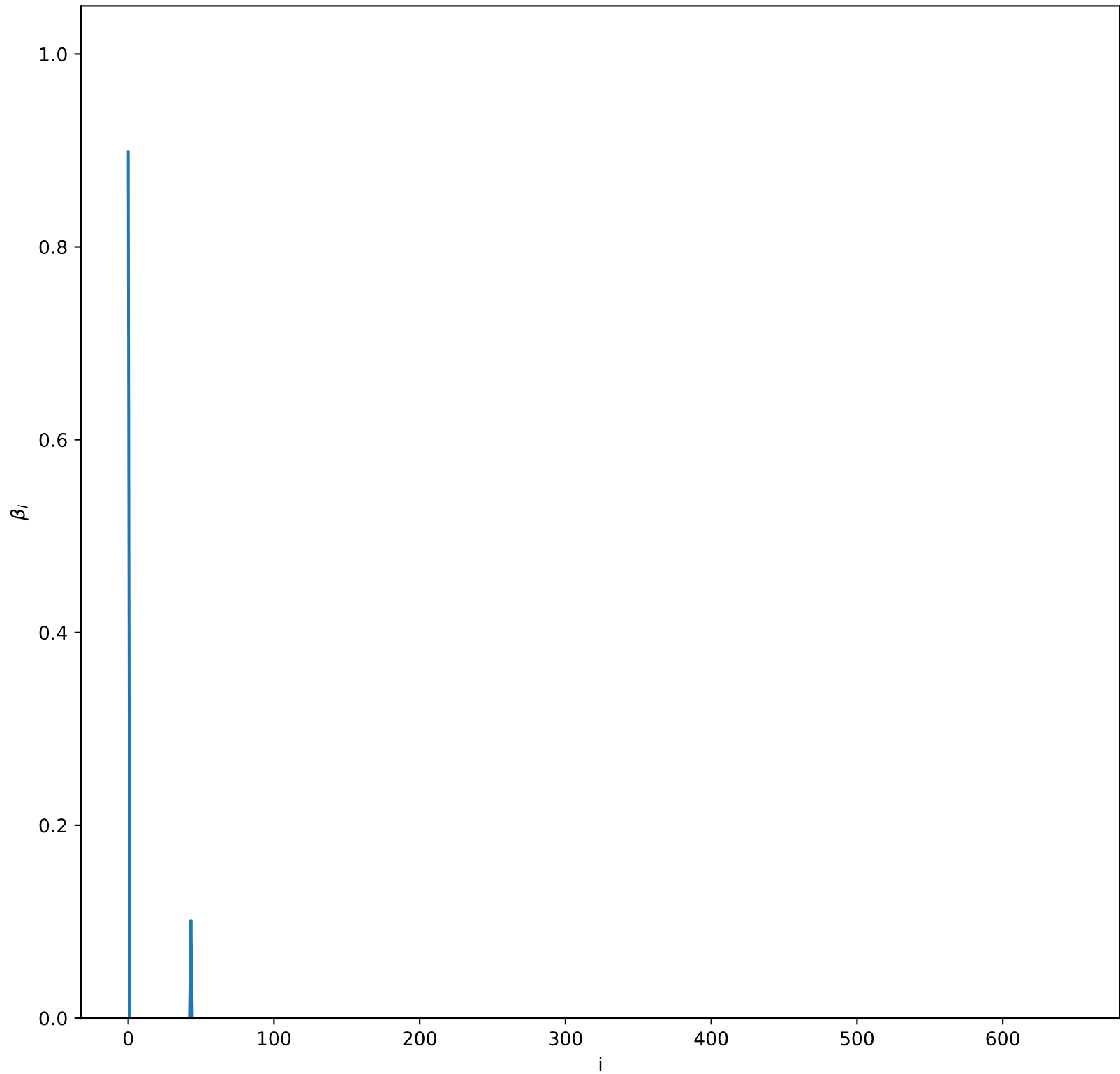
$\mu = 2.42$



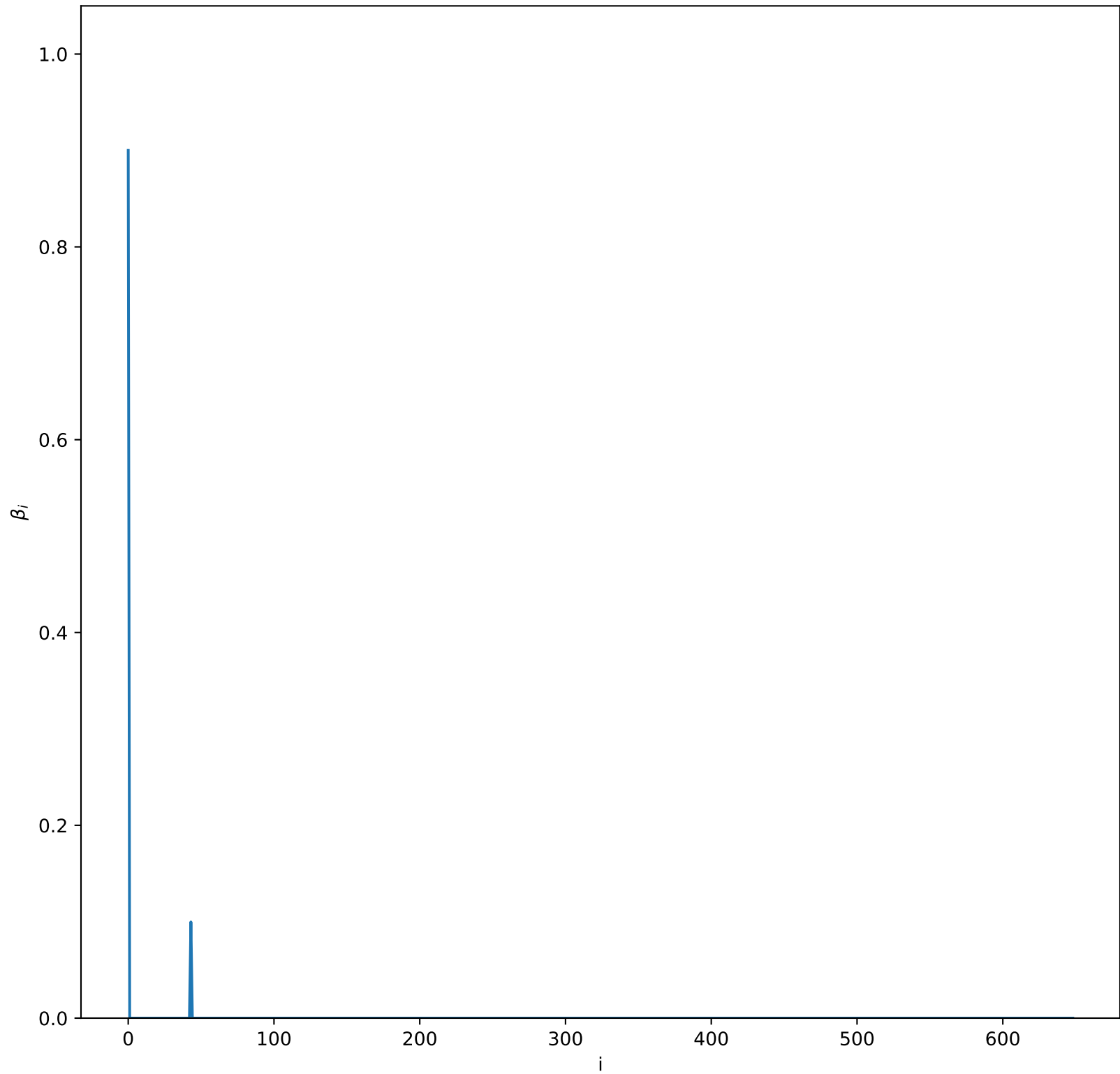
$\mu = 2.43$



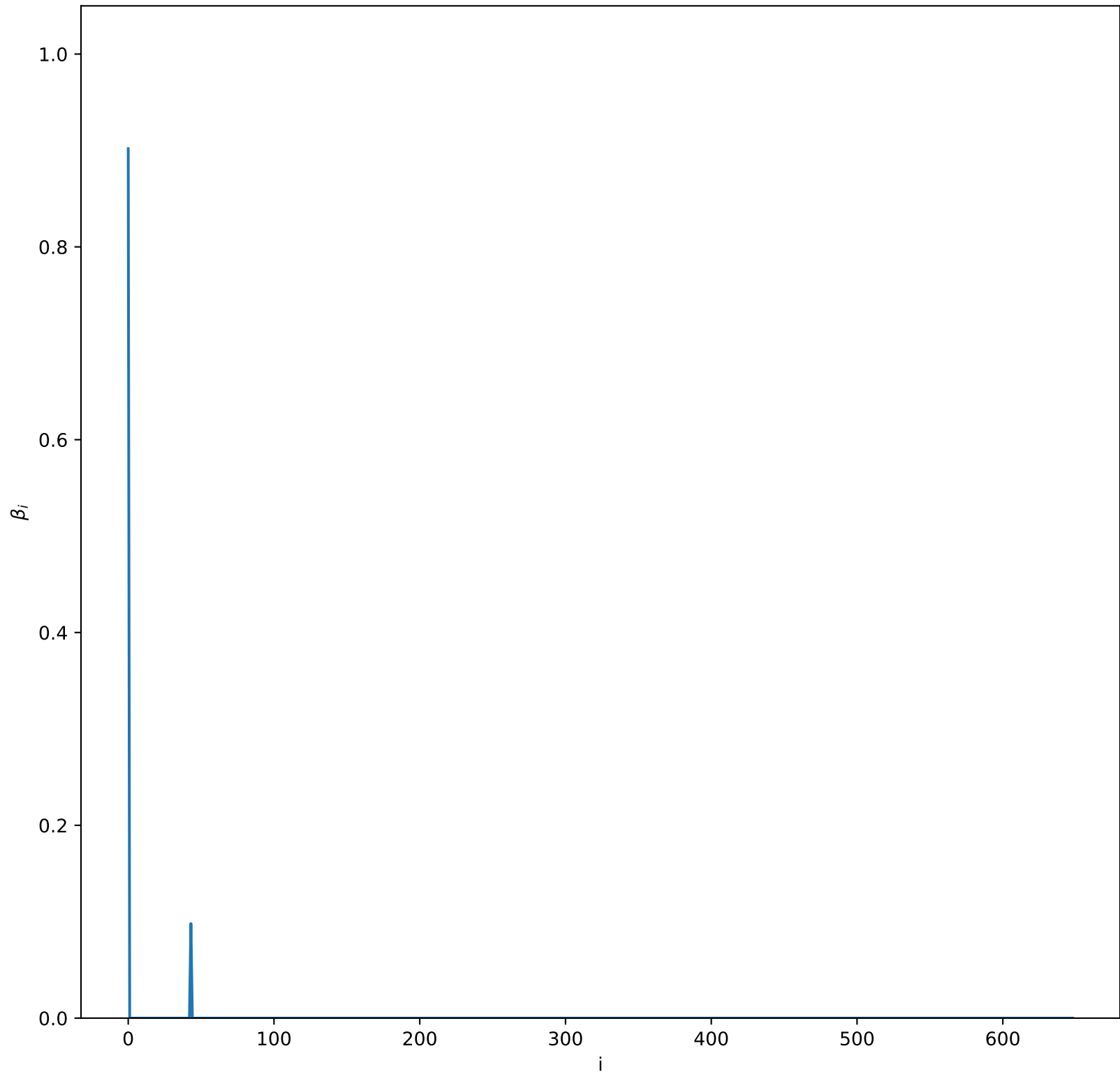
$\mu = 2.44$



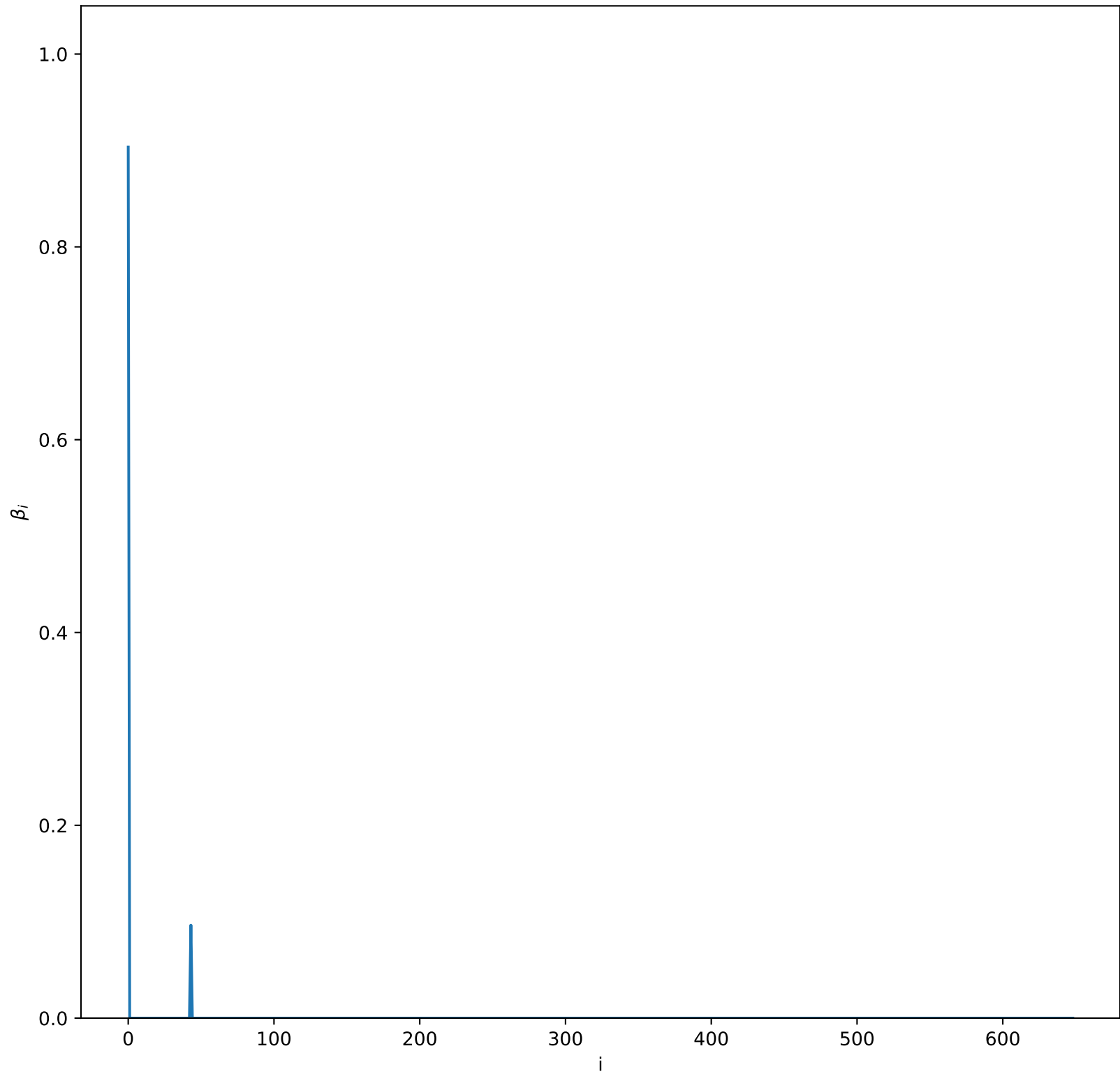
$\mu = 2.45$



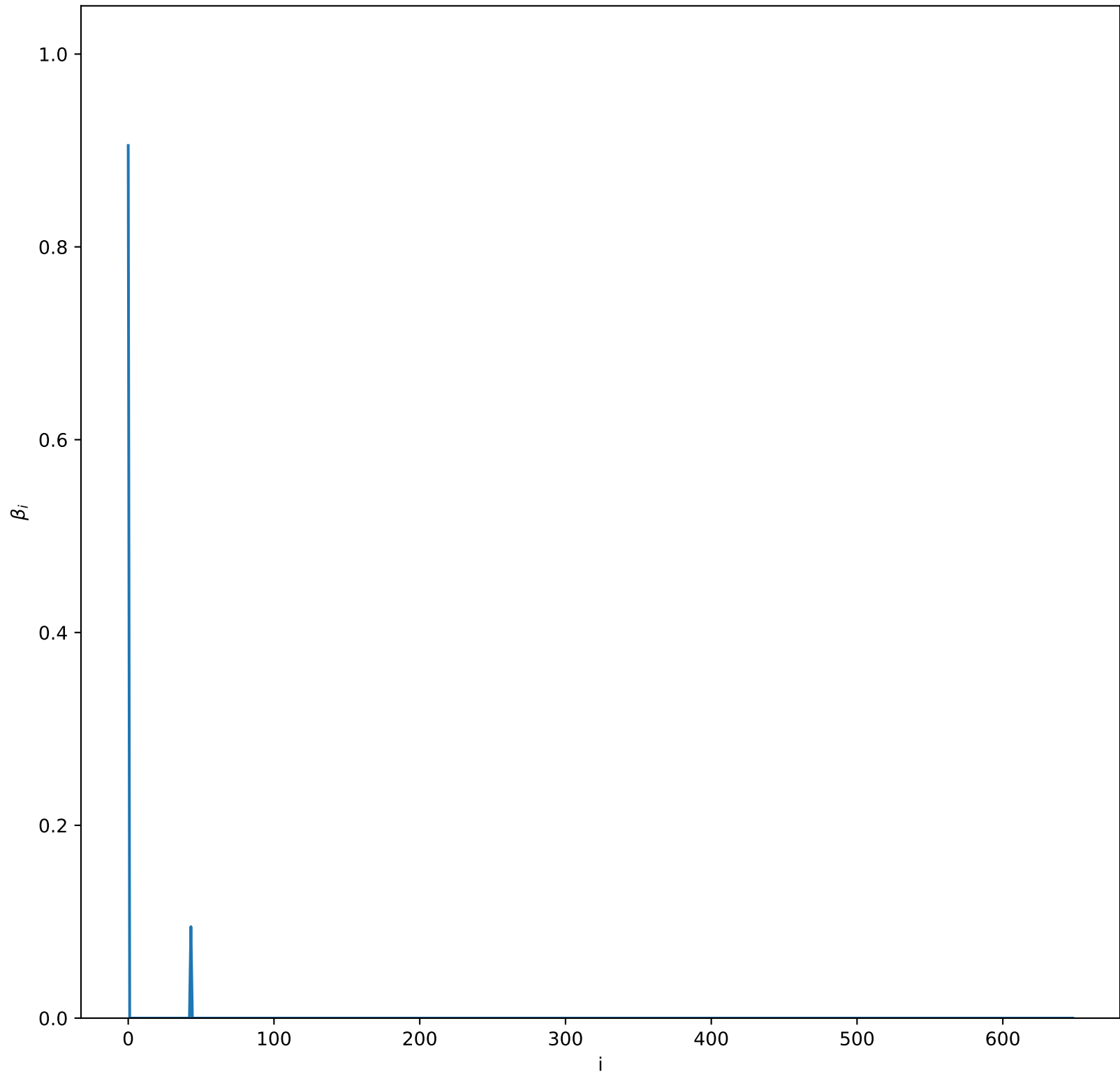
$\mu = 2.46$



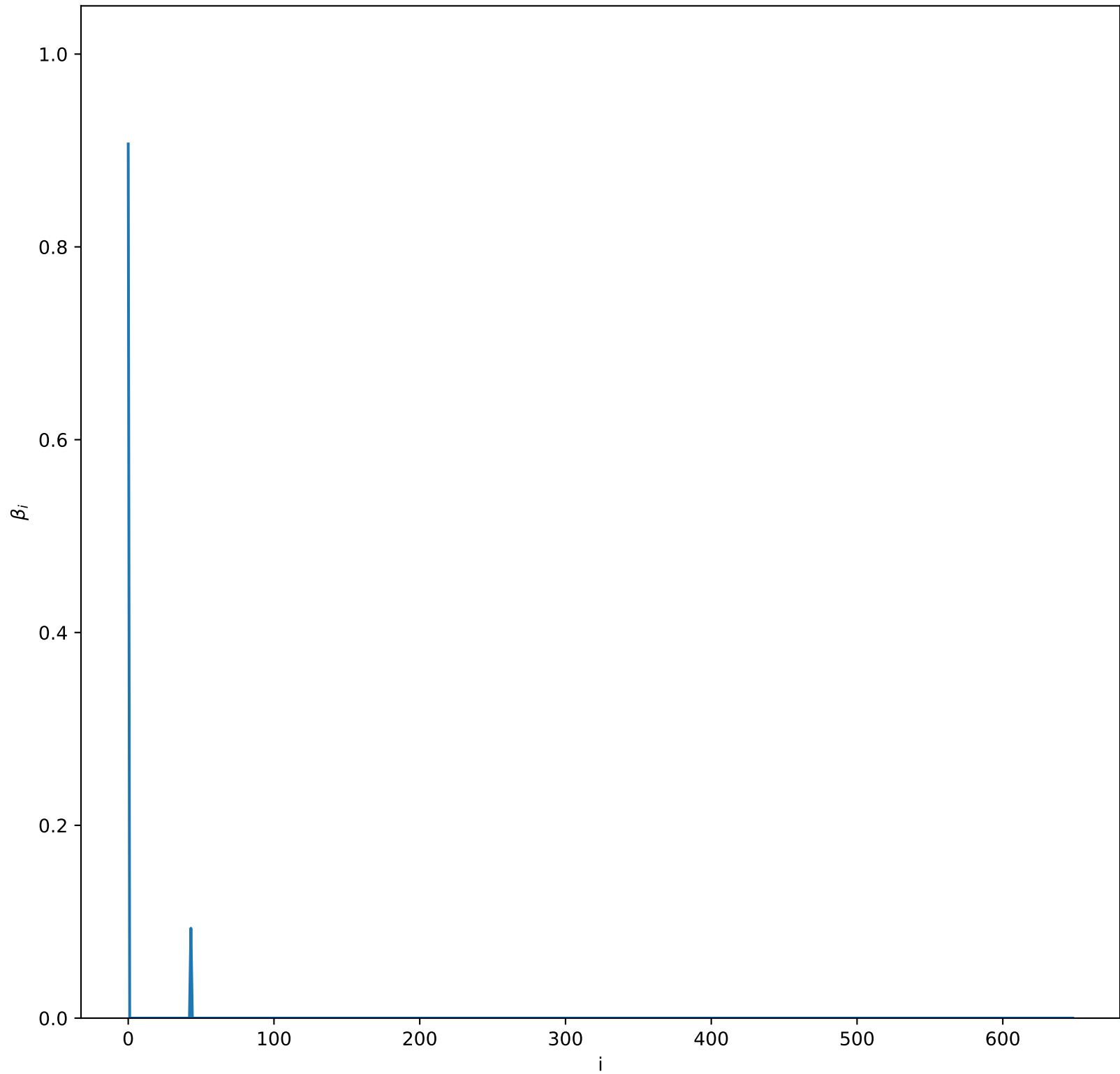
$\mu = 2.47$



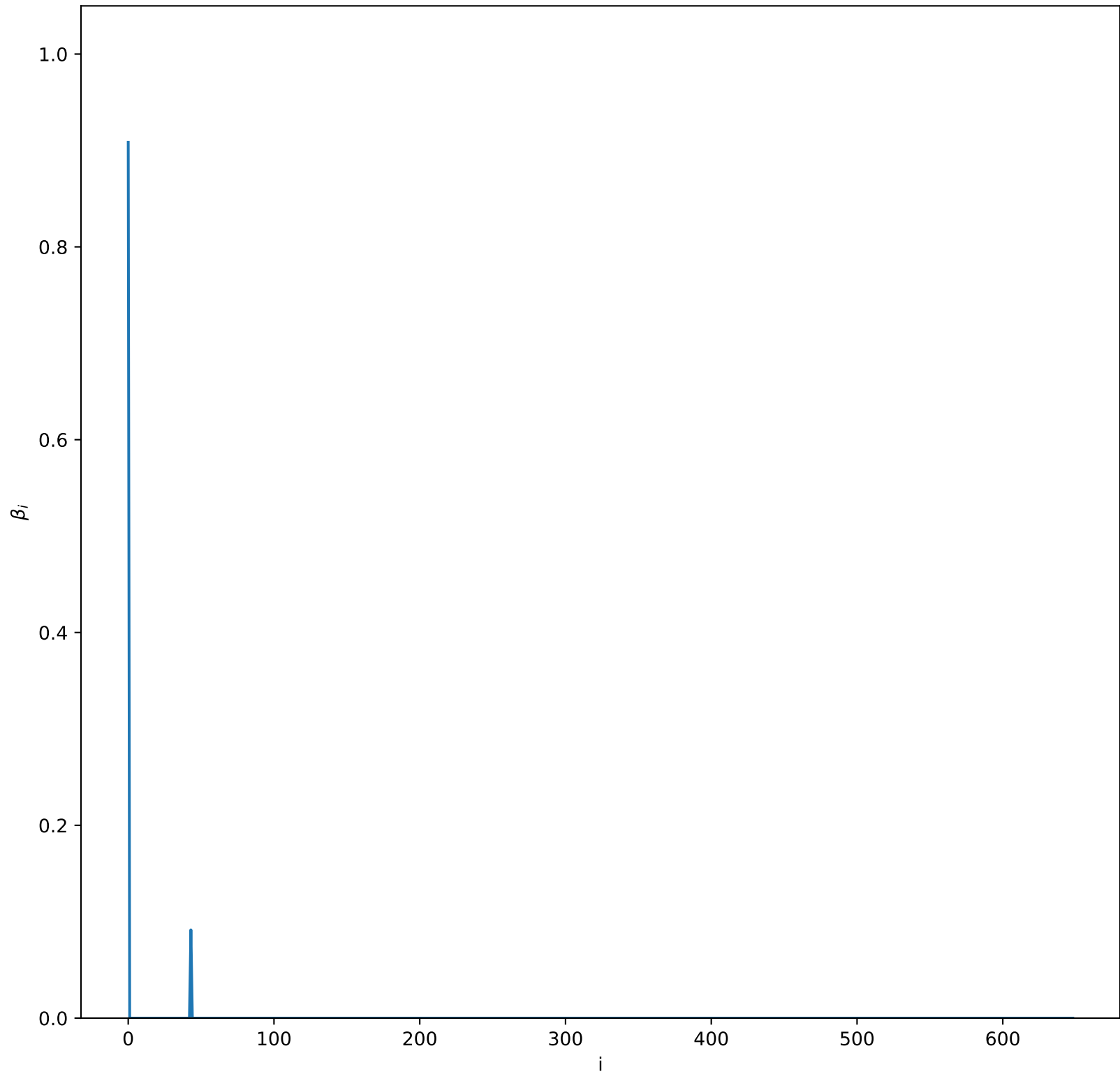
$\mu = 2.48$



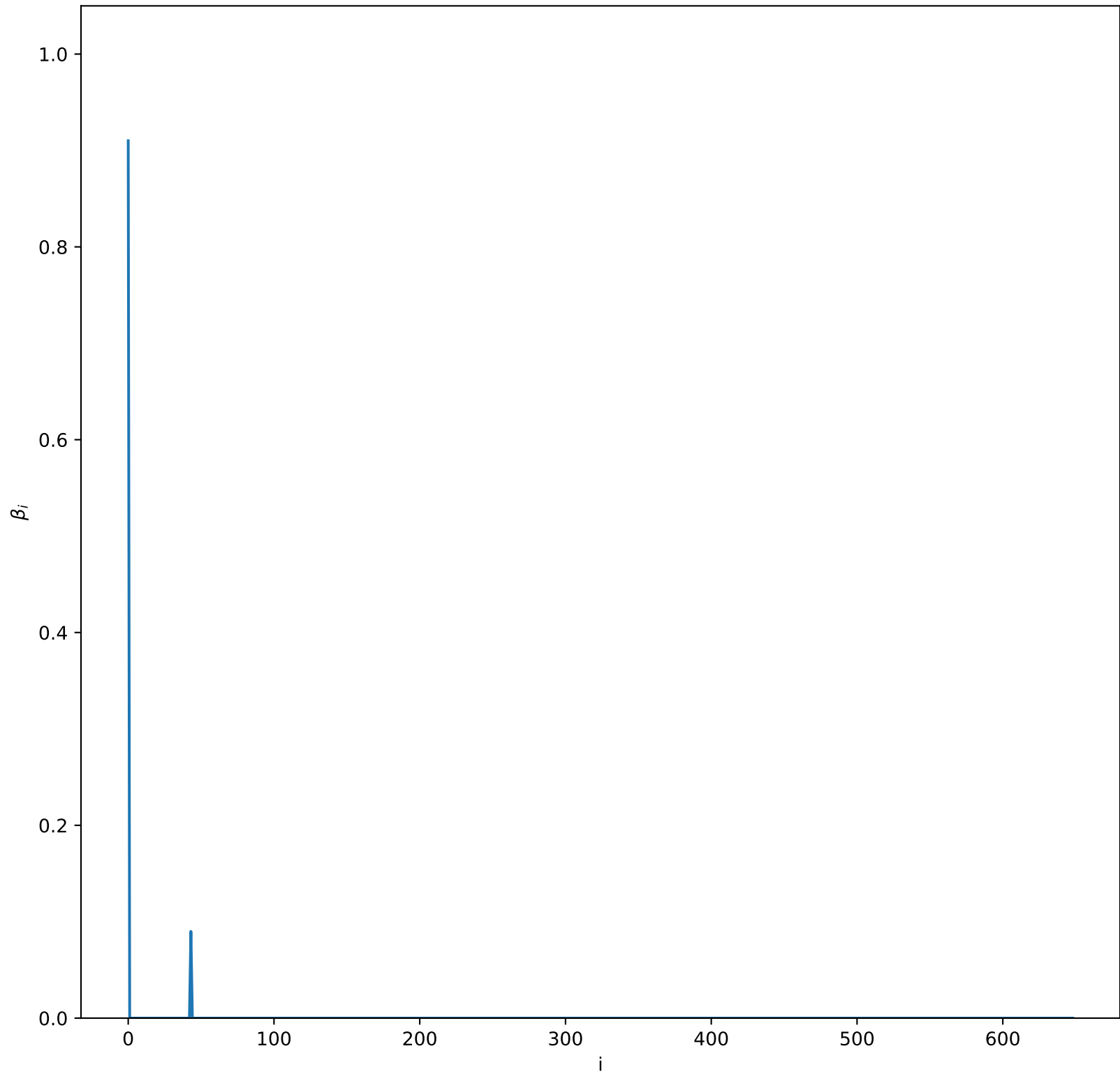
$\mu = 2.49$



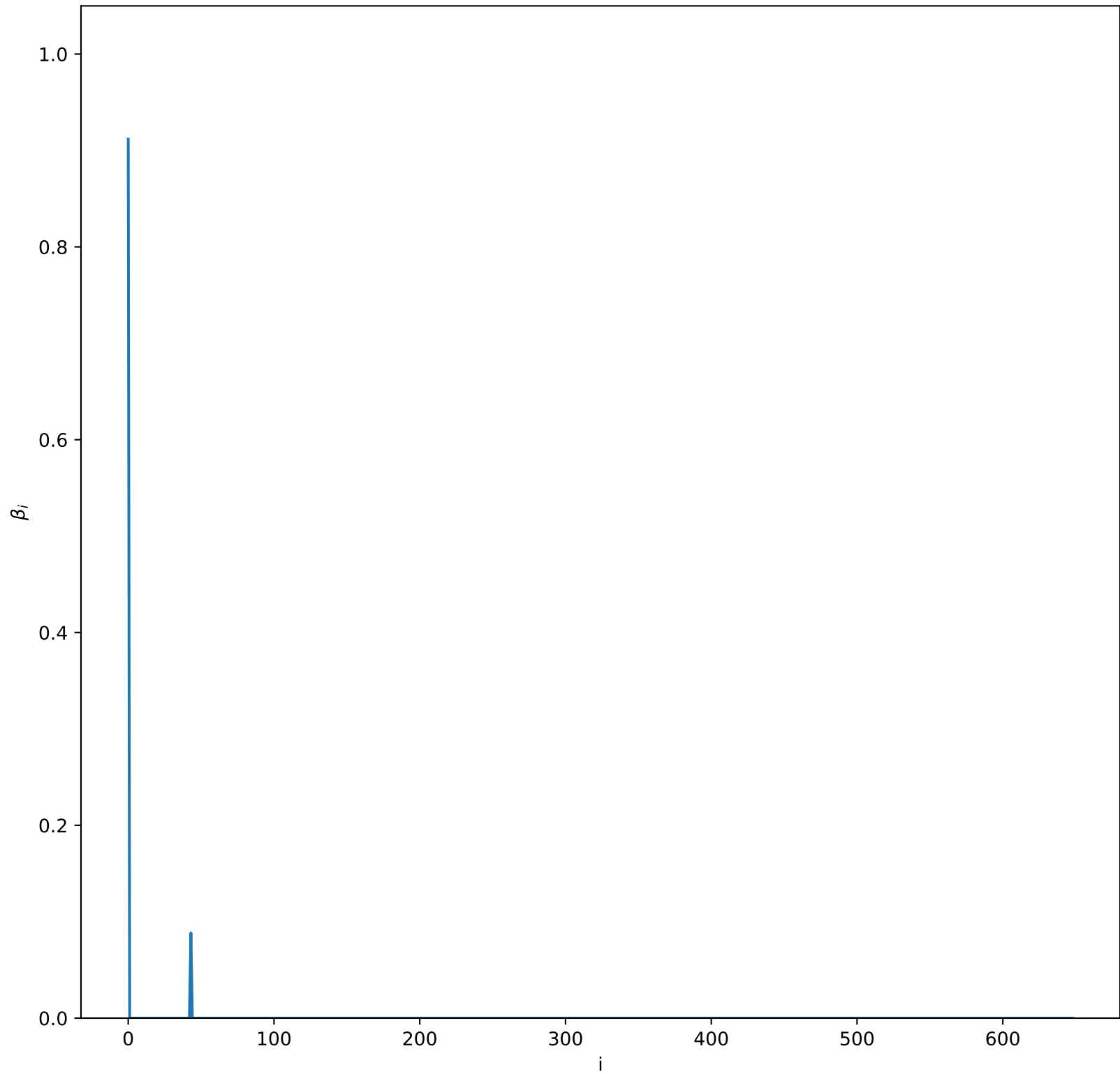
$\mu = 2.50$



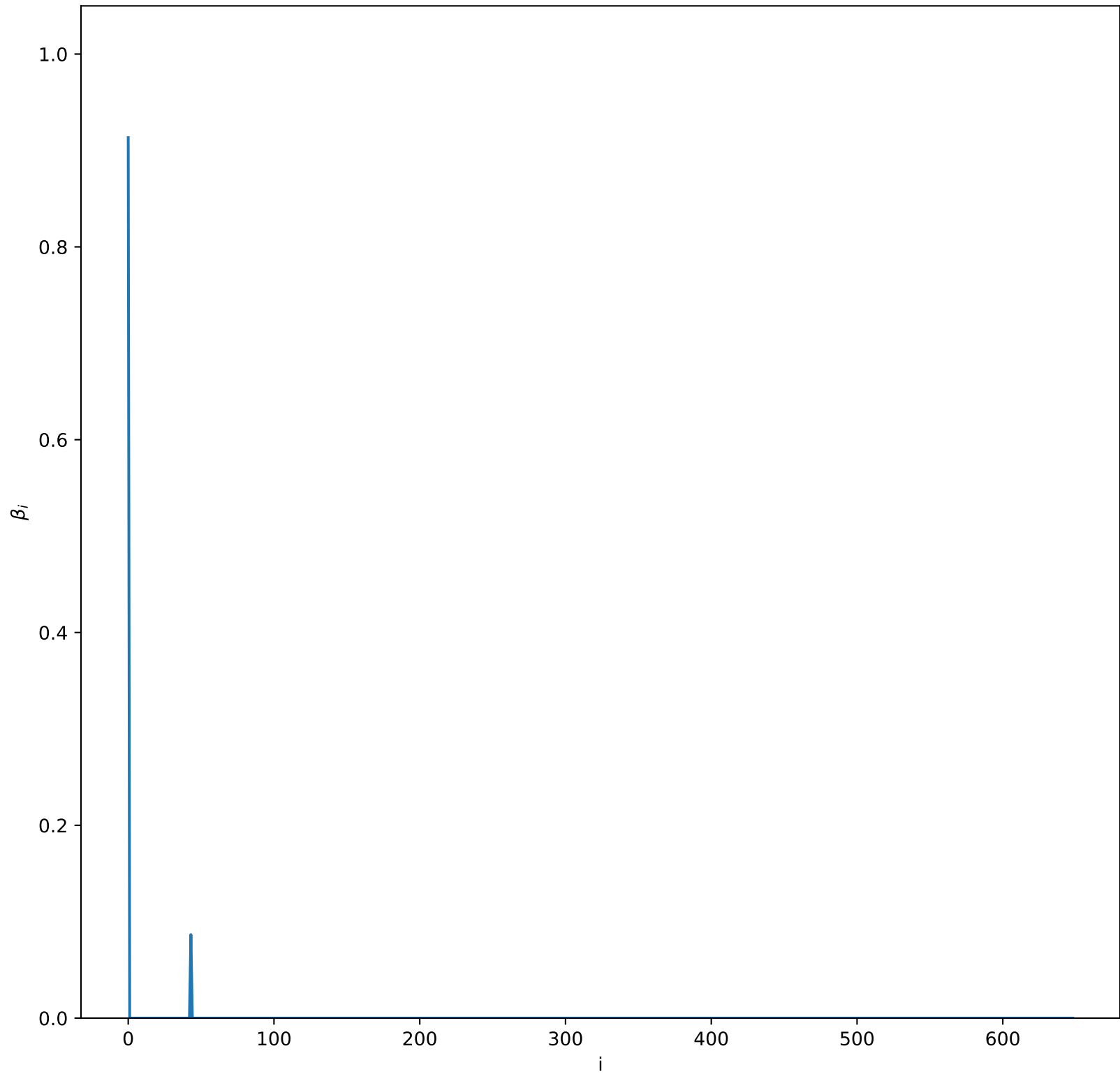
$\mu = 2.51$



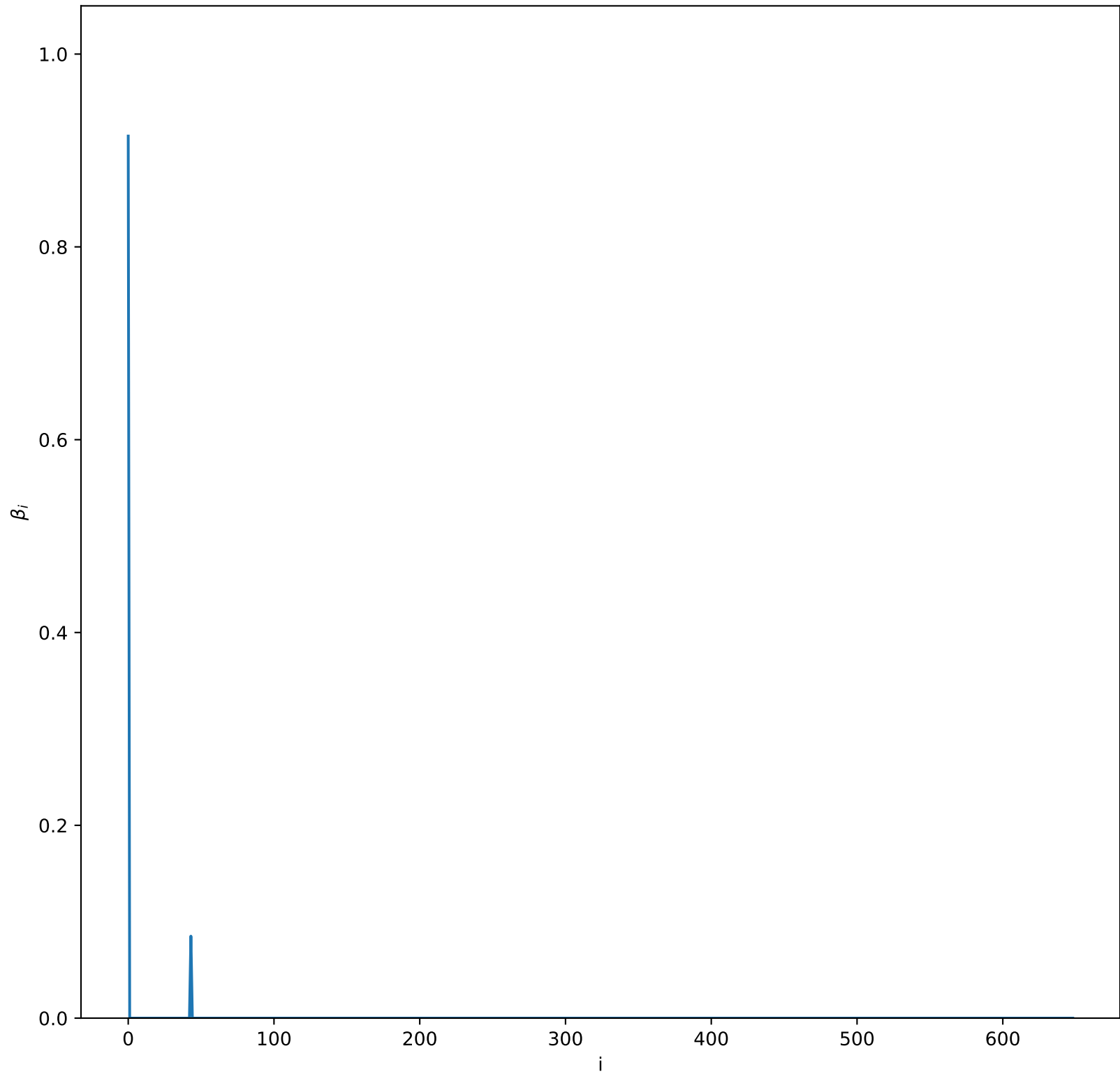
$\mu = 2.52$



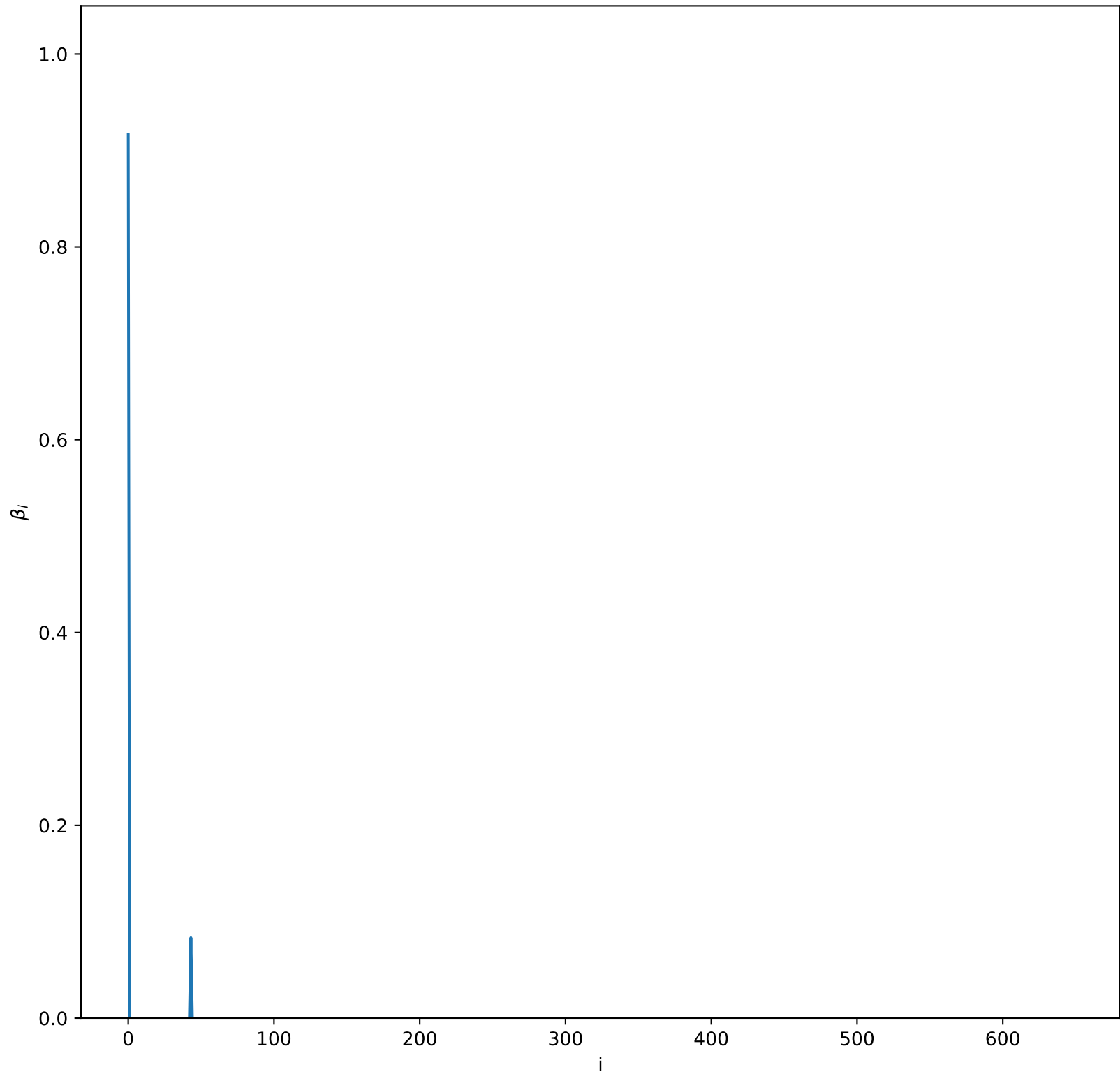
$\mu = 2.53$



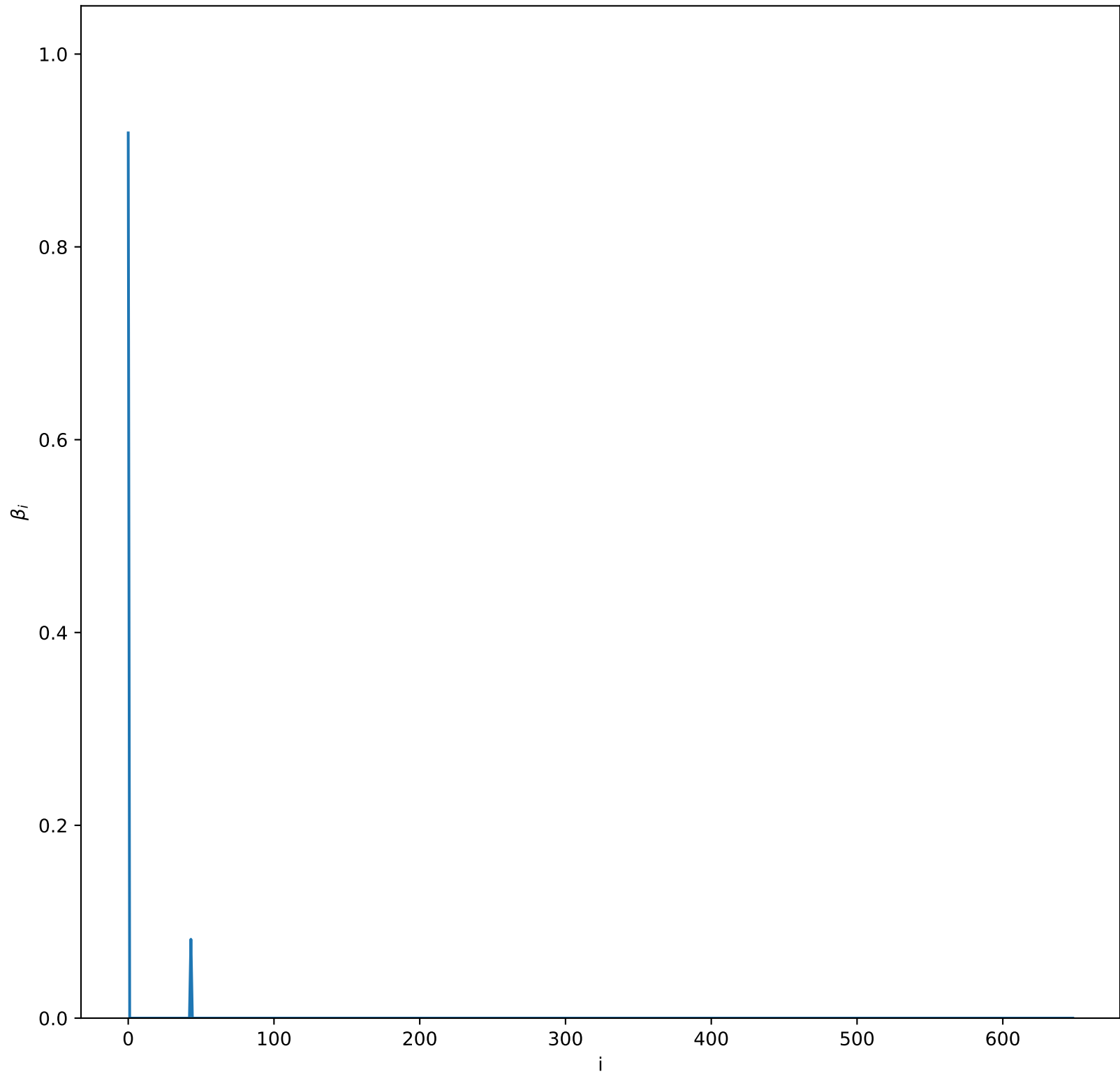
$\mu = 2.54$



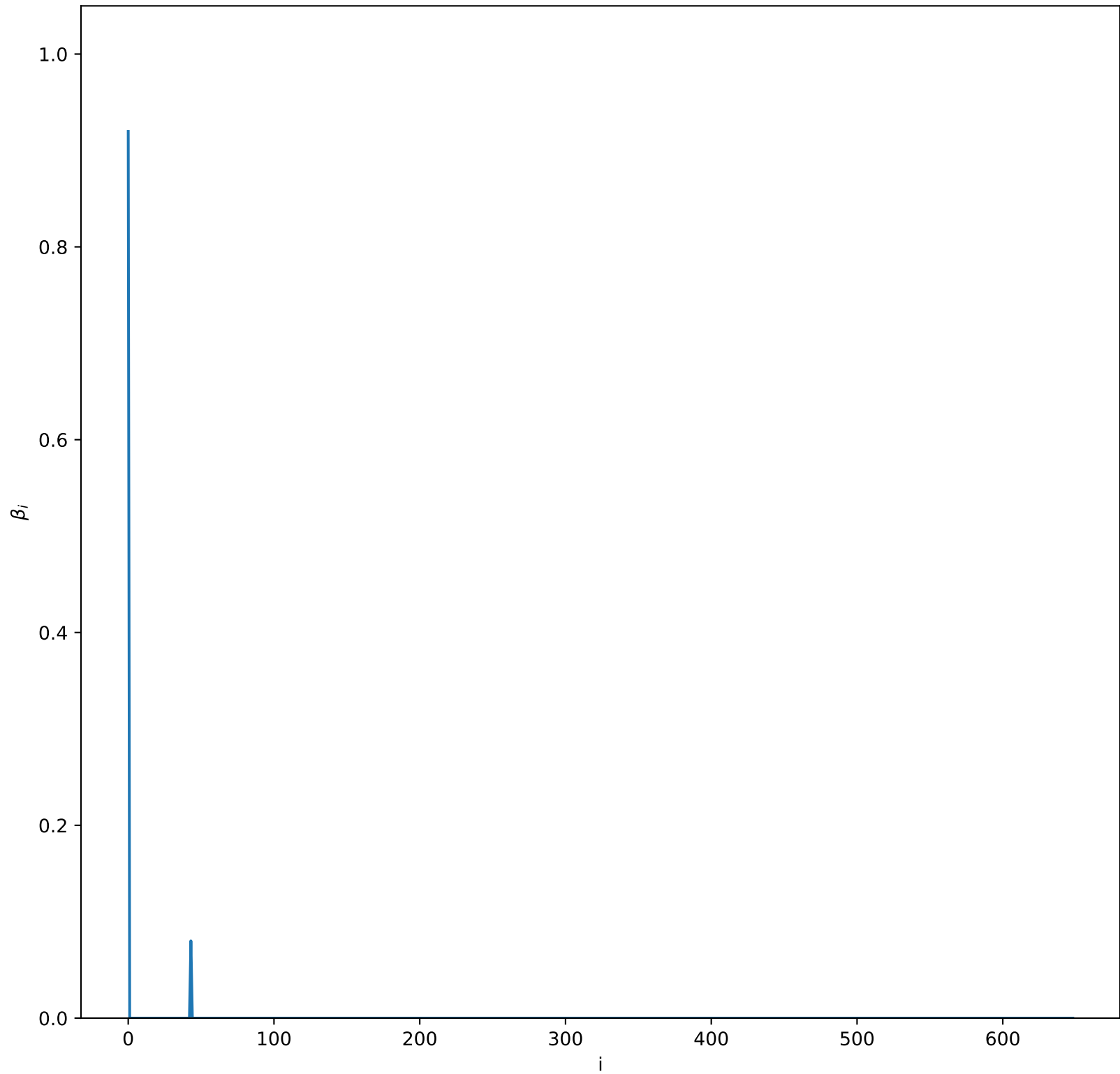
$\mu = 2.55$



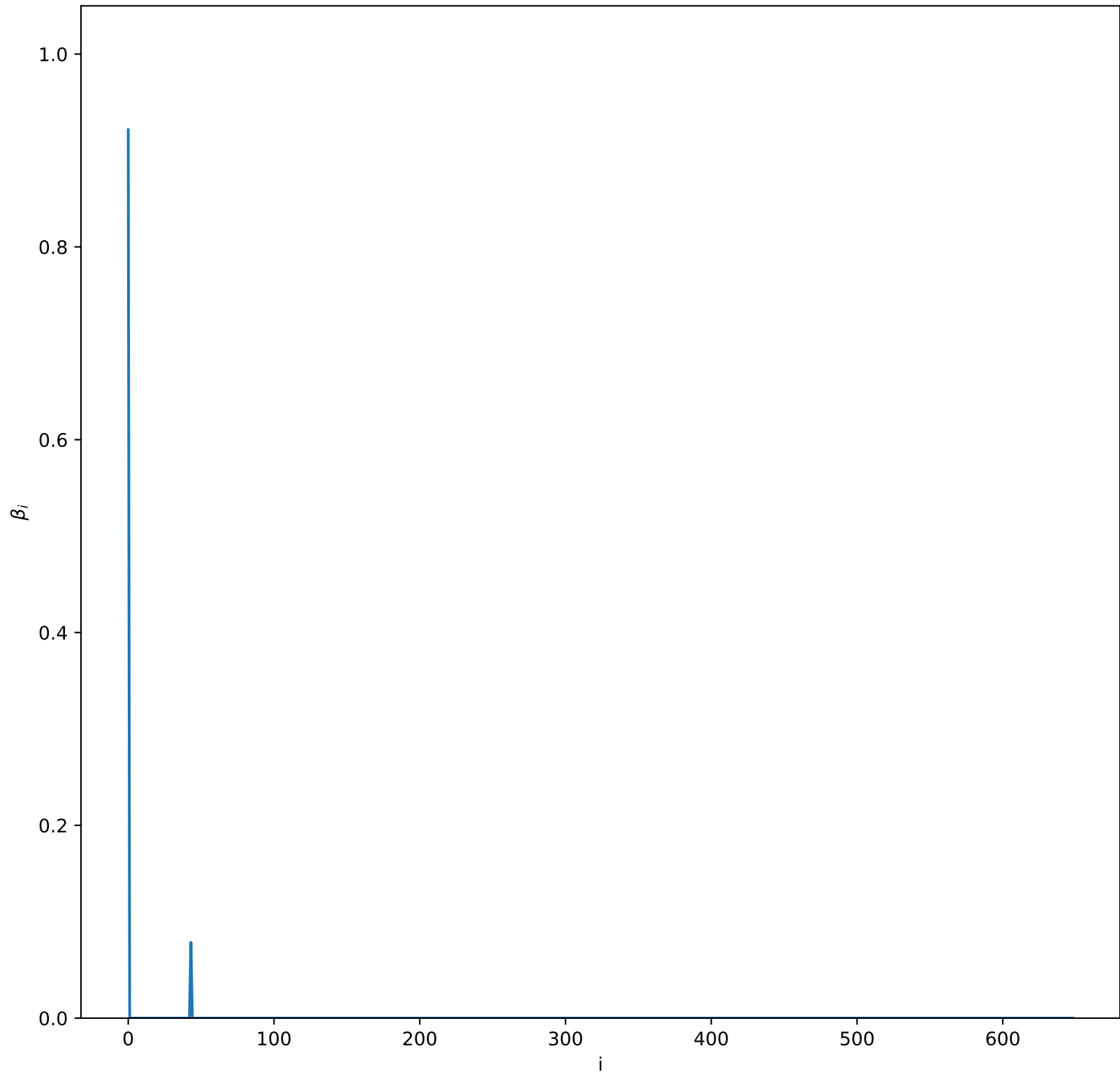
$\mu = 2.56$



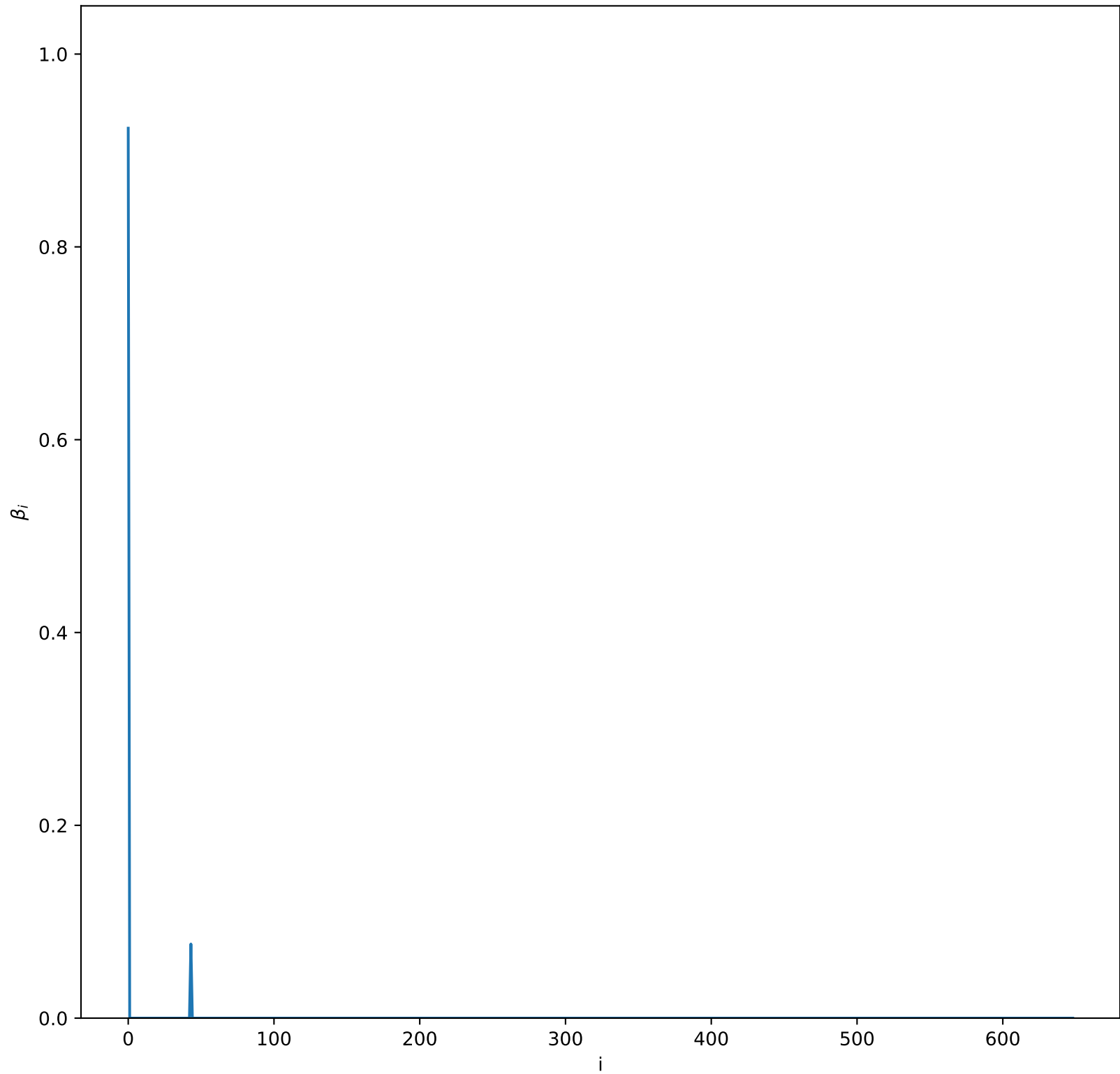
$\mu = 2.57$



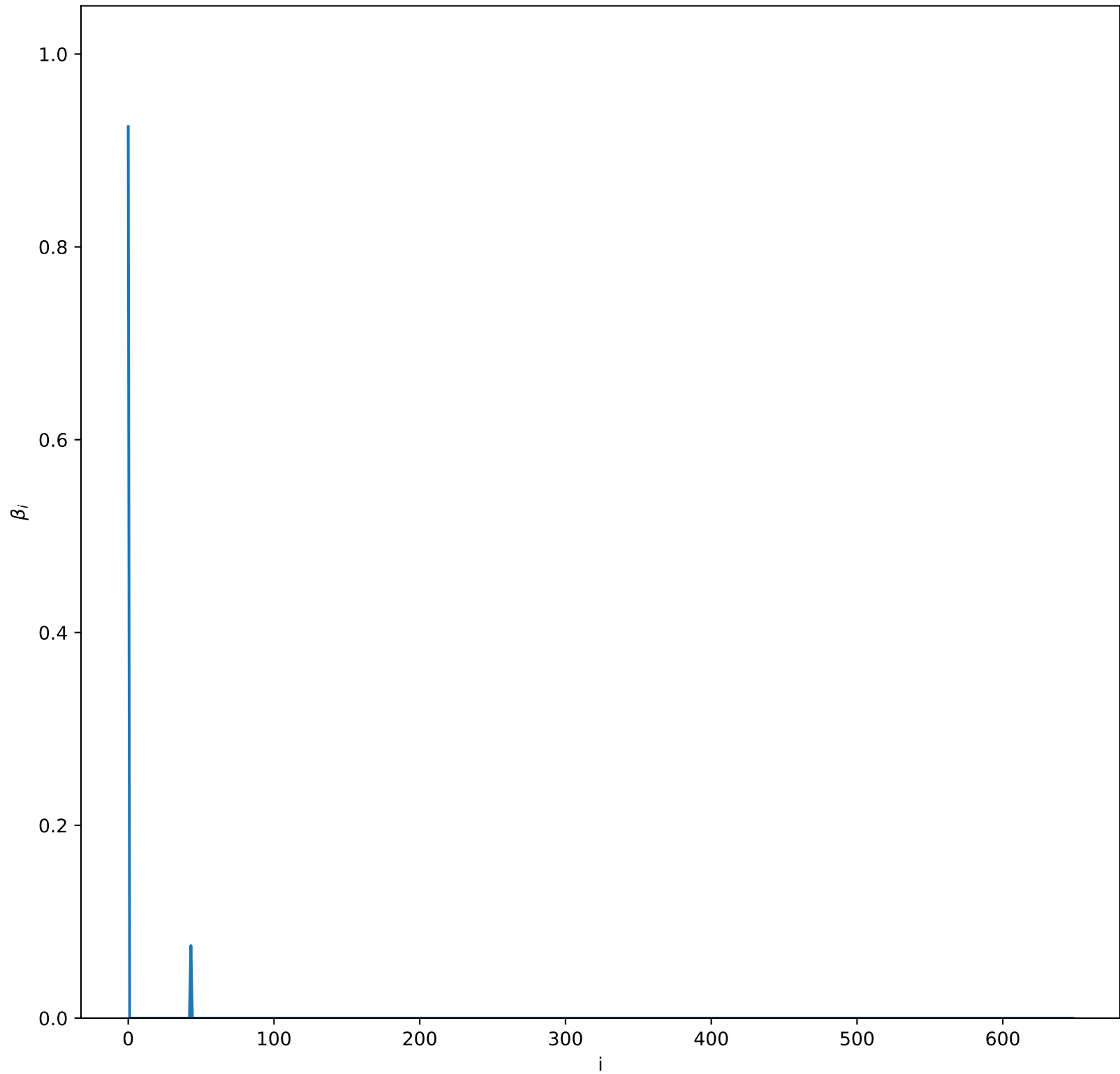
$\mu = 2.58$



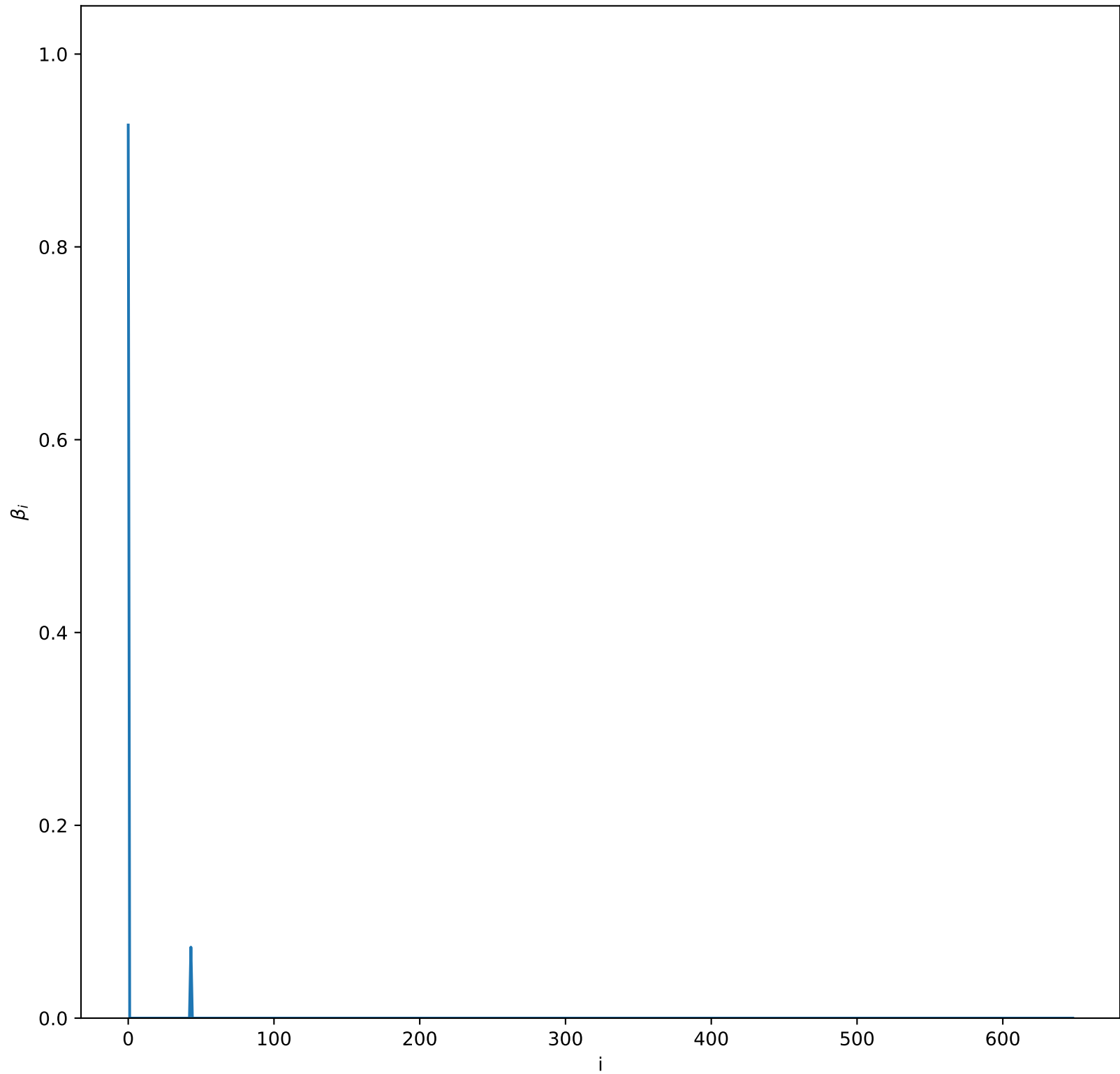
$\mu = 2.59$



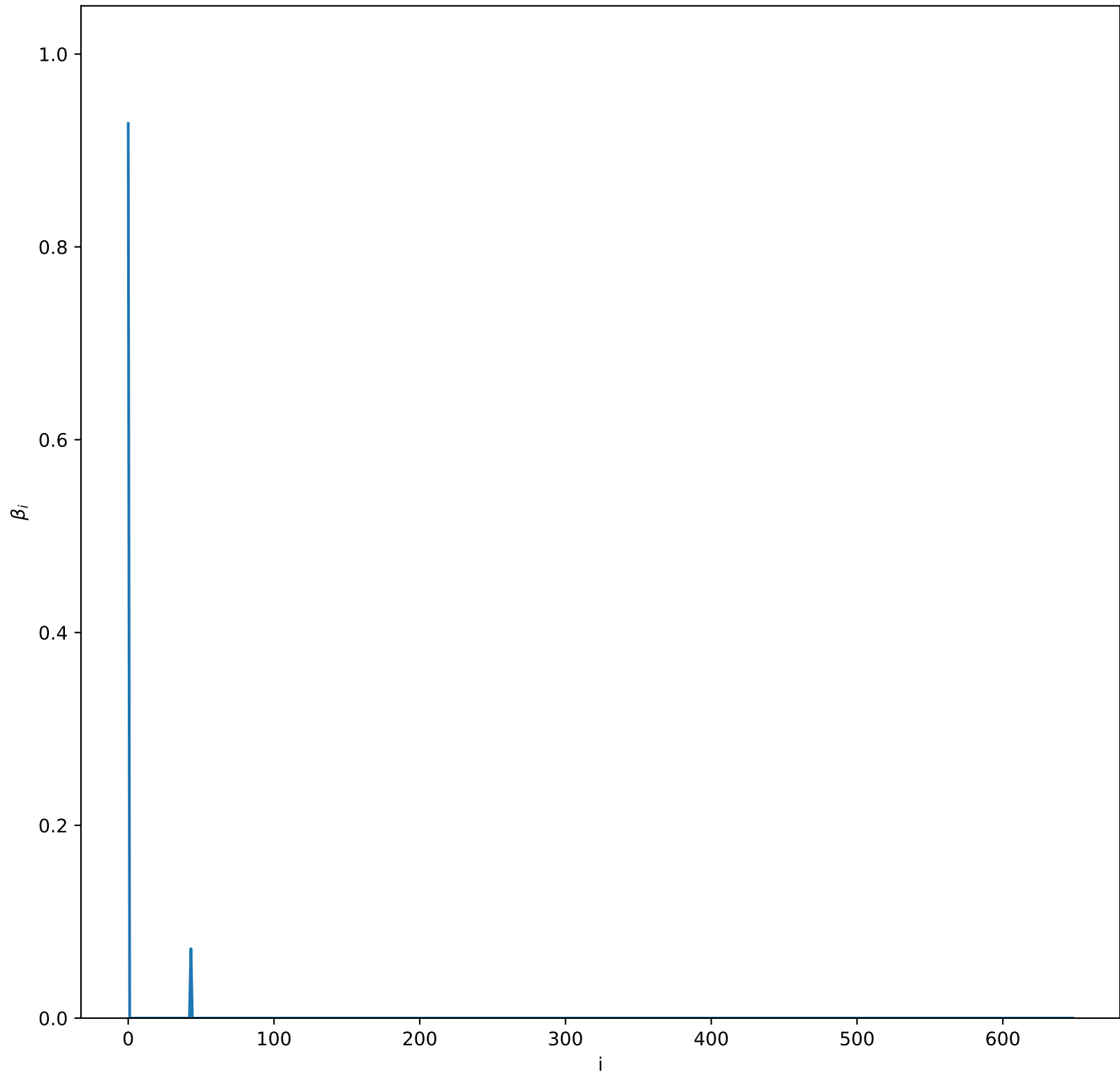
$\mu = 2.60$



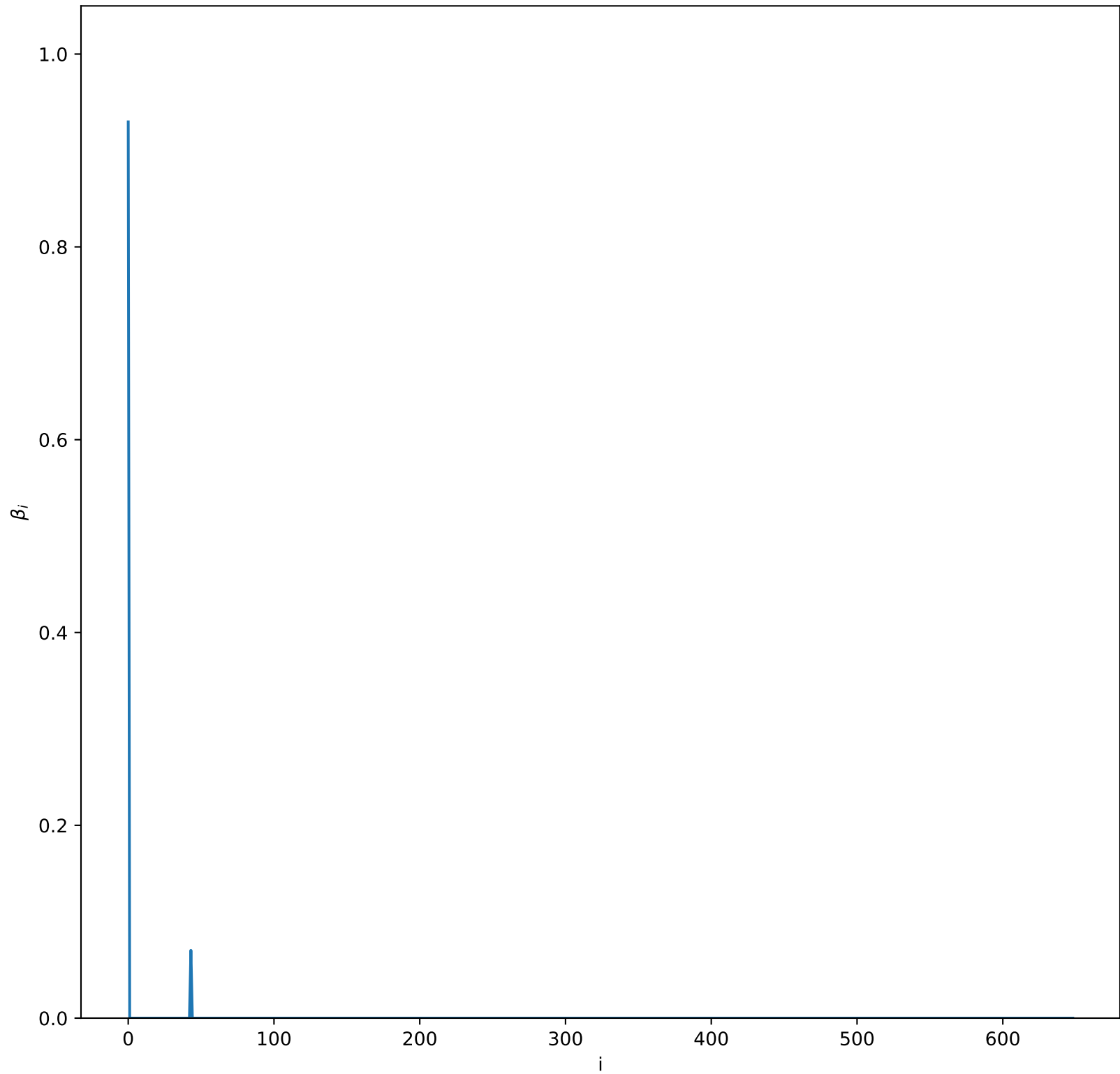
$\mu = 2.61$



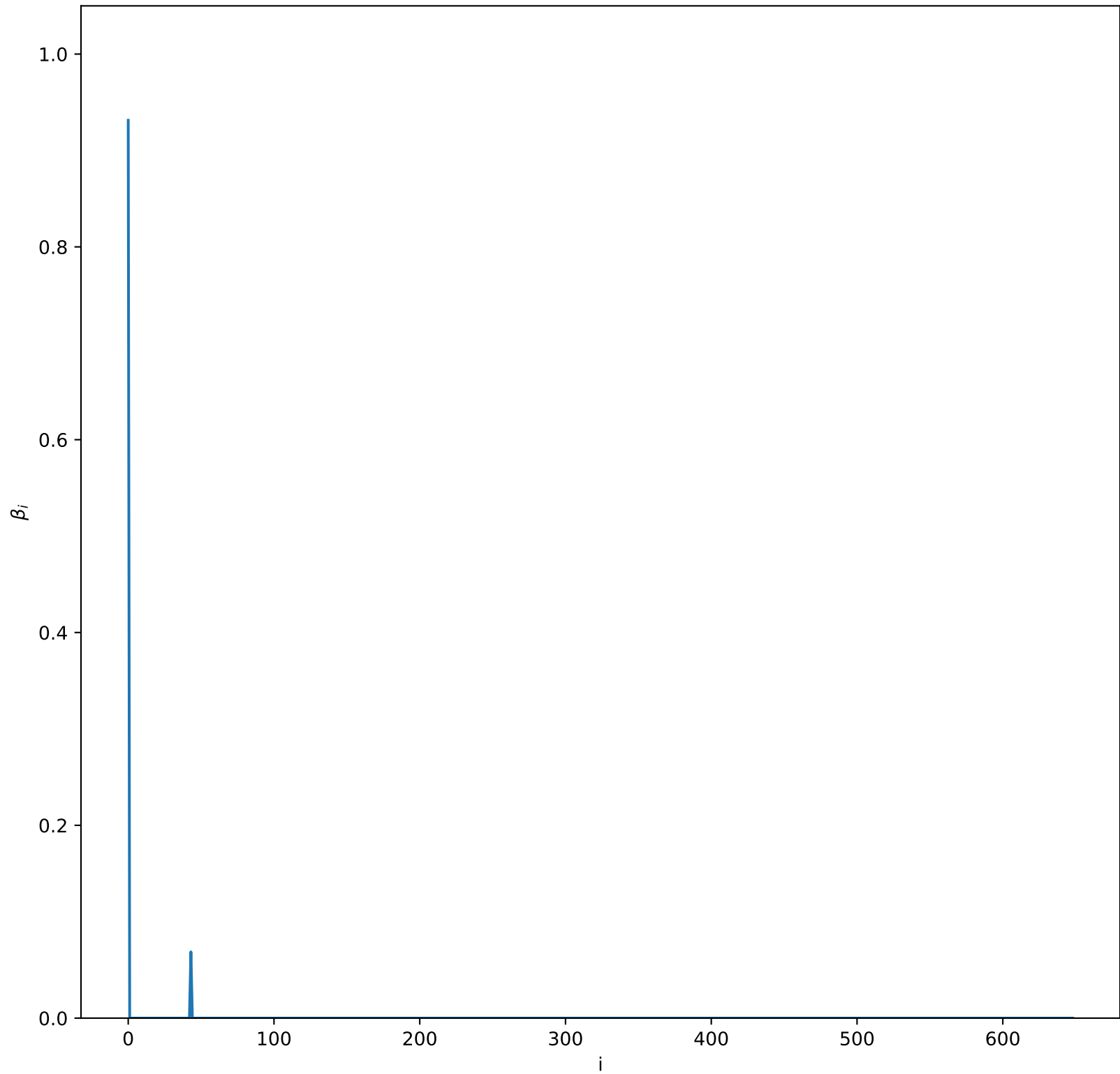
$\mu = 2.62$



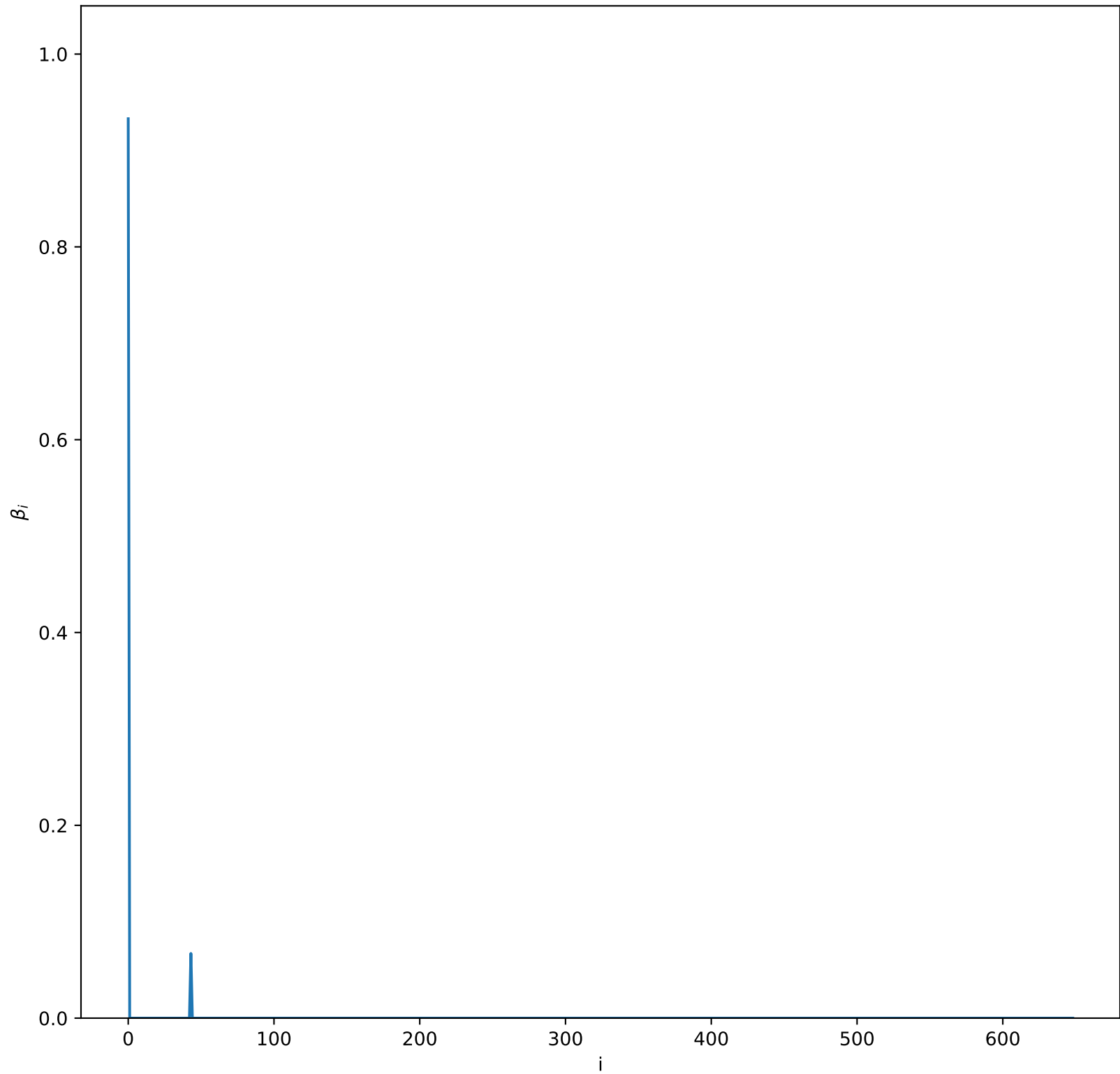
$\mu = 2.63$



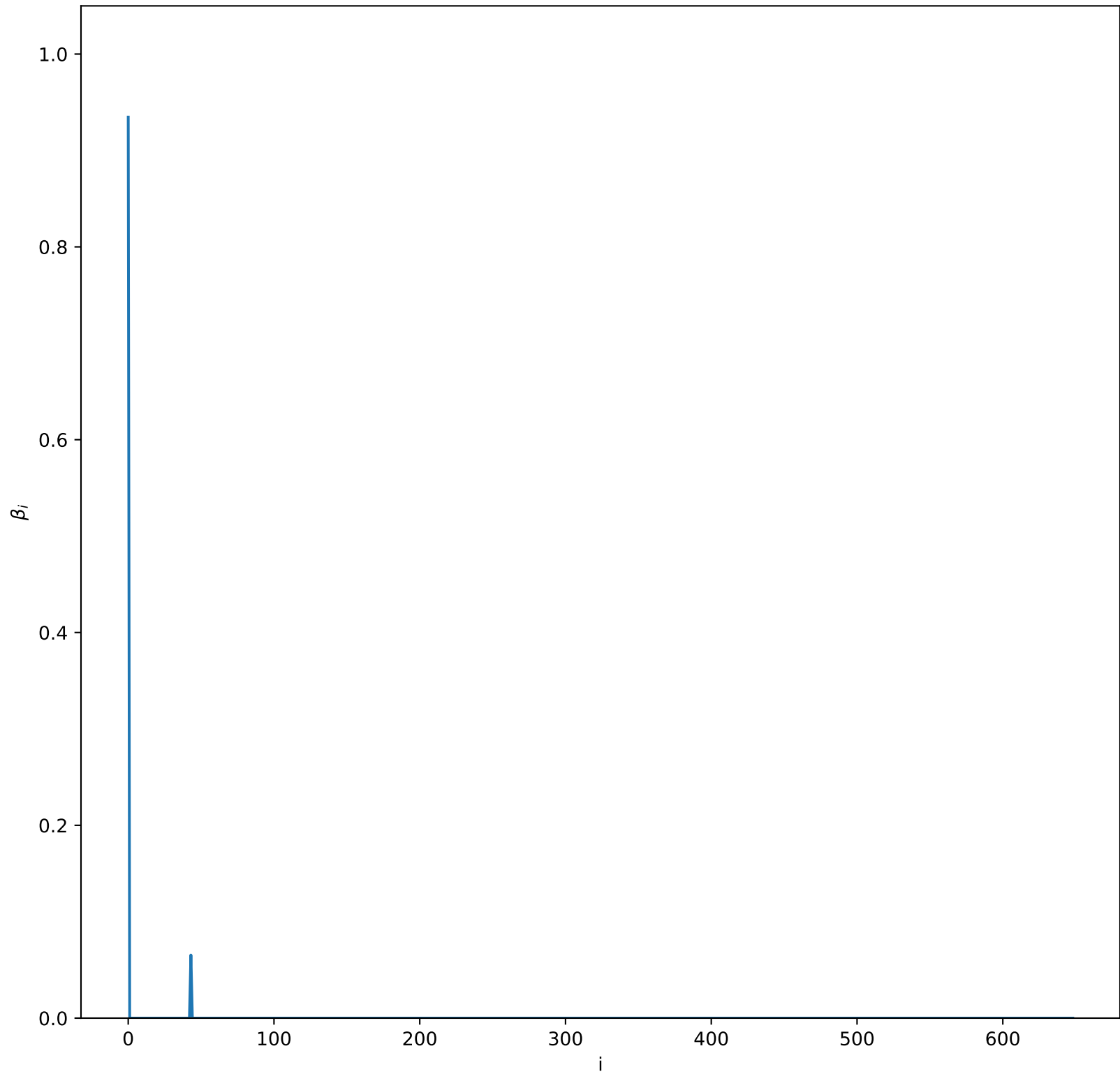
$\mu = 2.64$



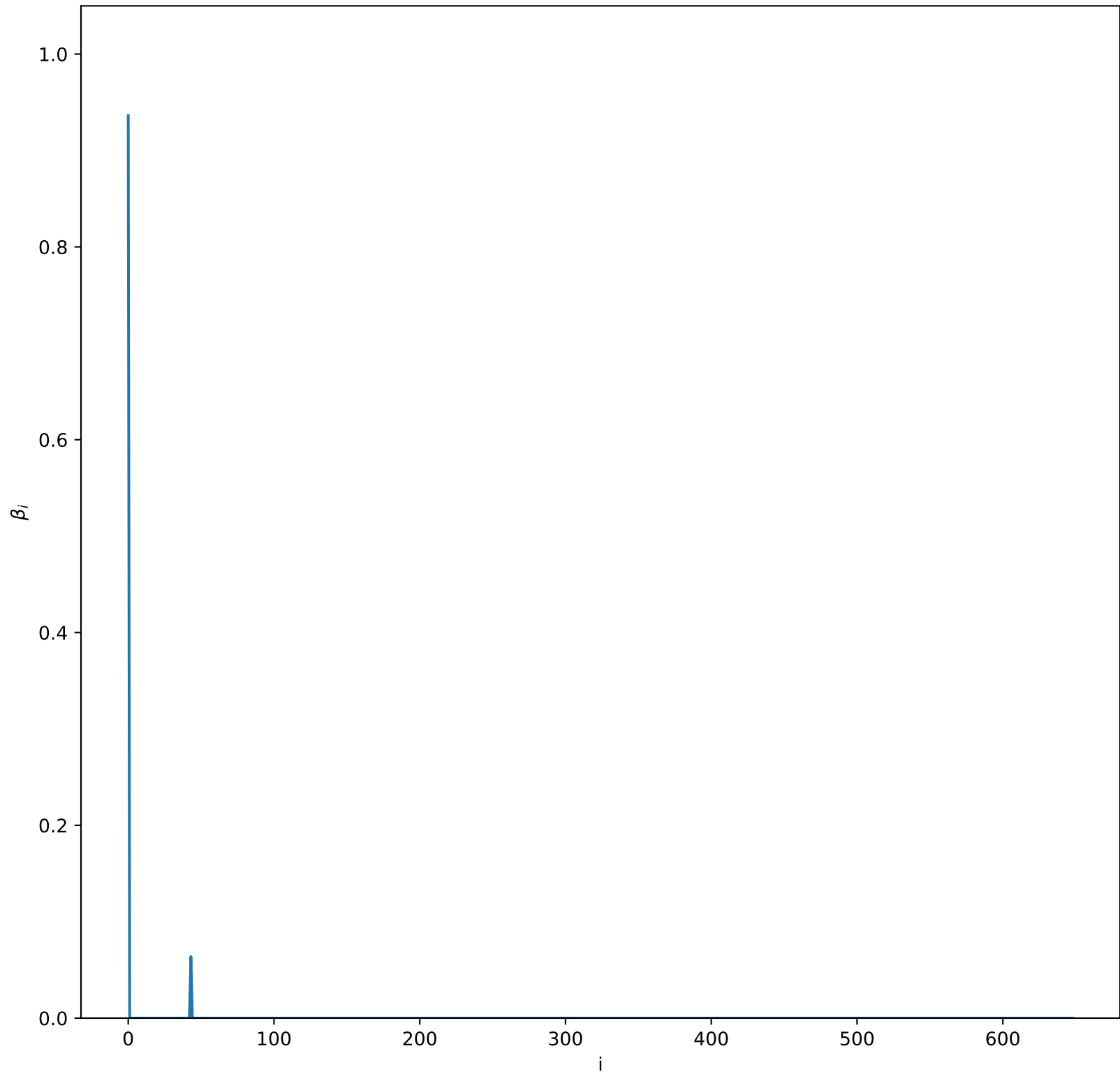
$\mu = 2.65$



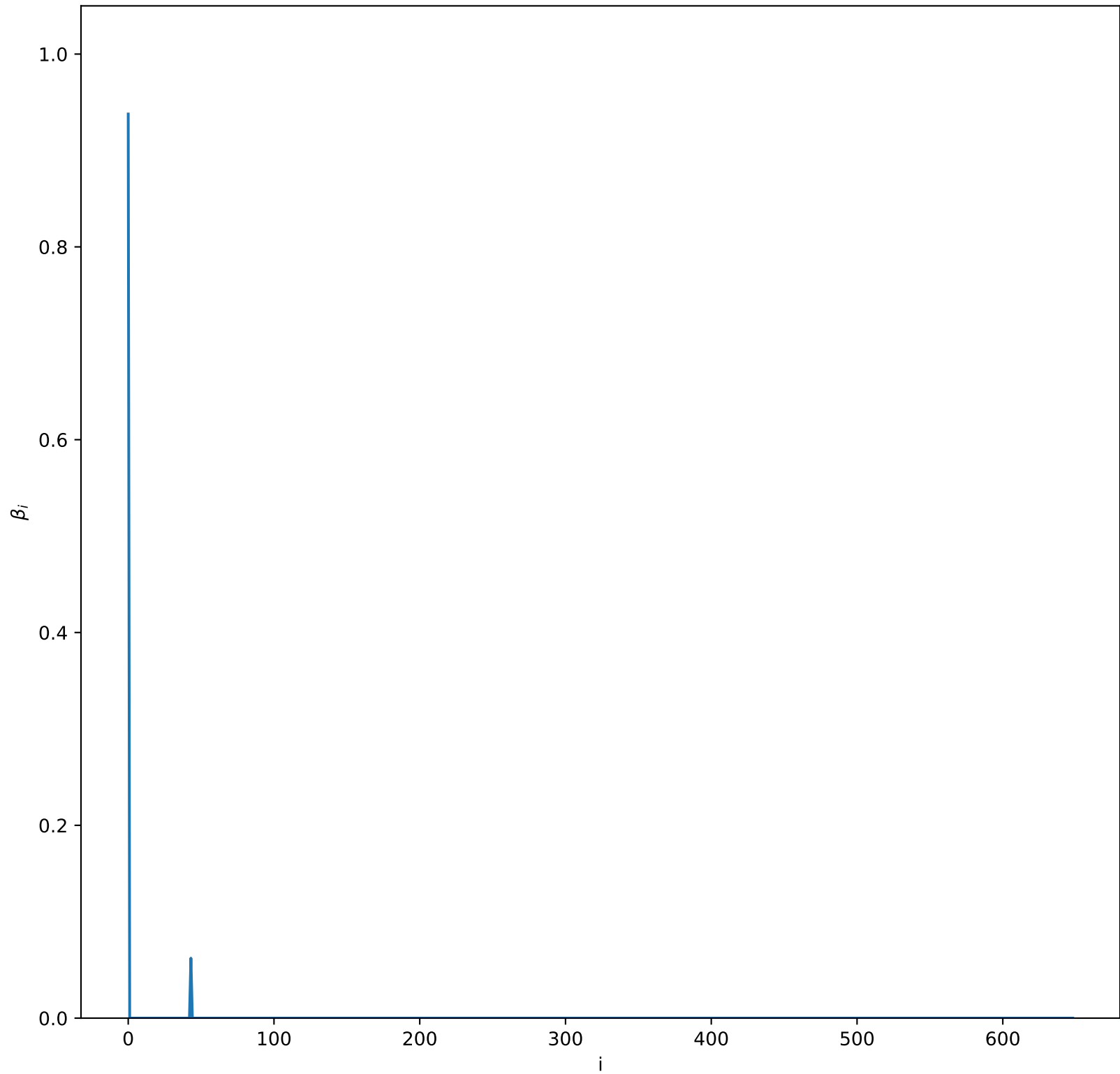
$\mu = 2.66$



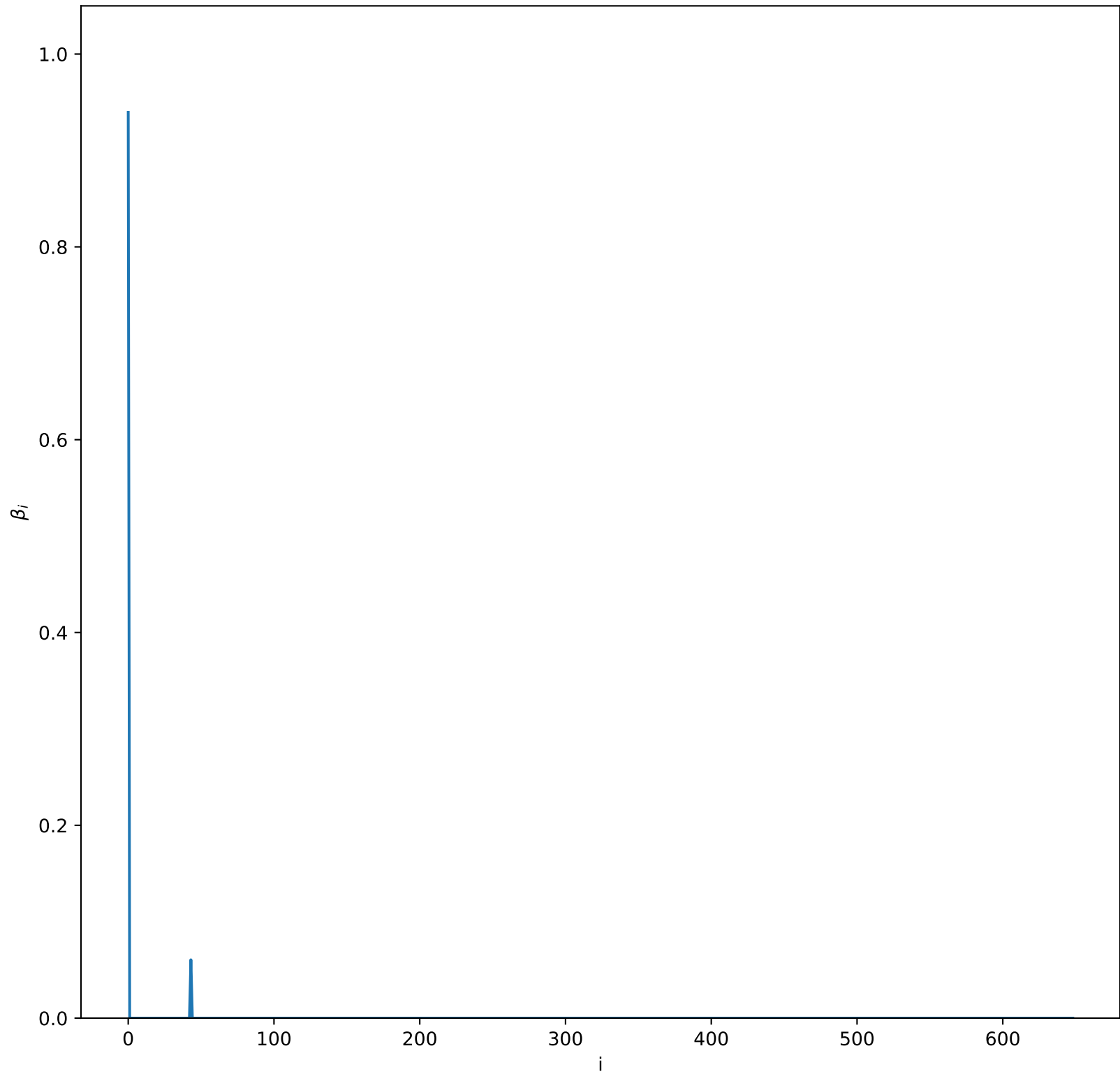
$\mu = 2.67$



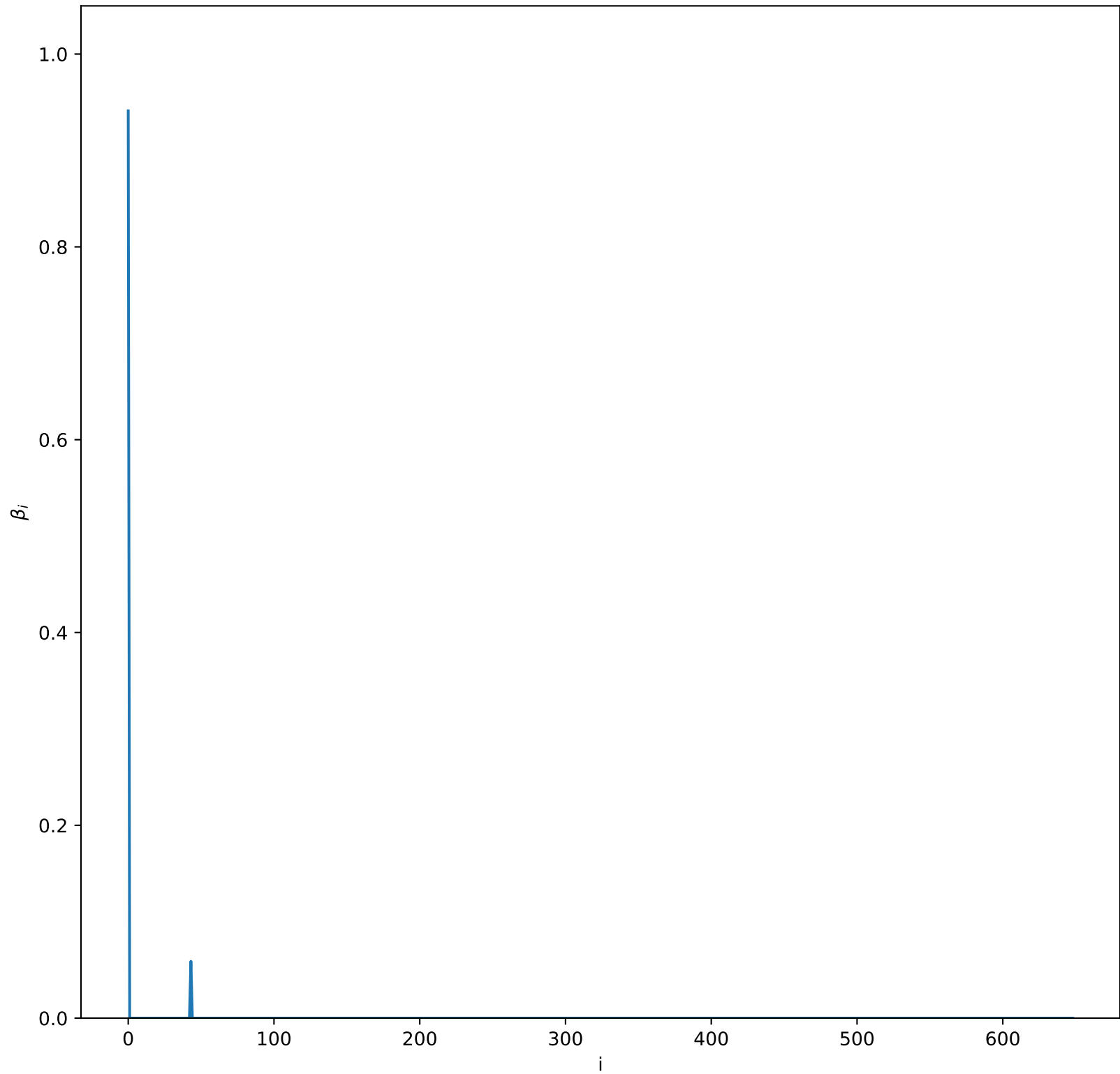
$\mu = 2.68$



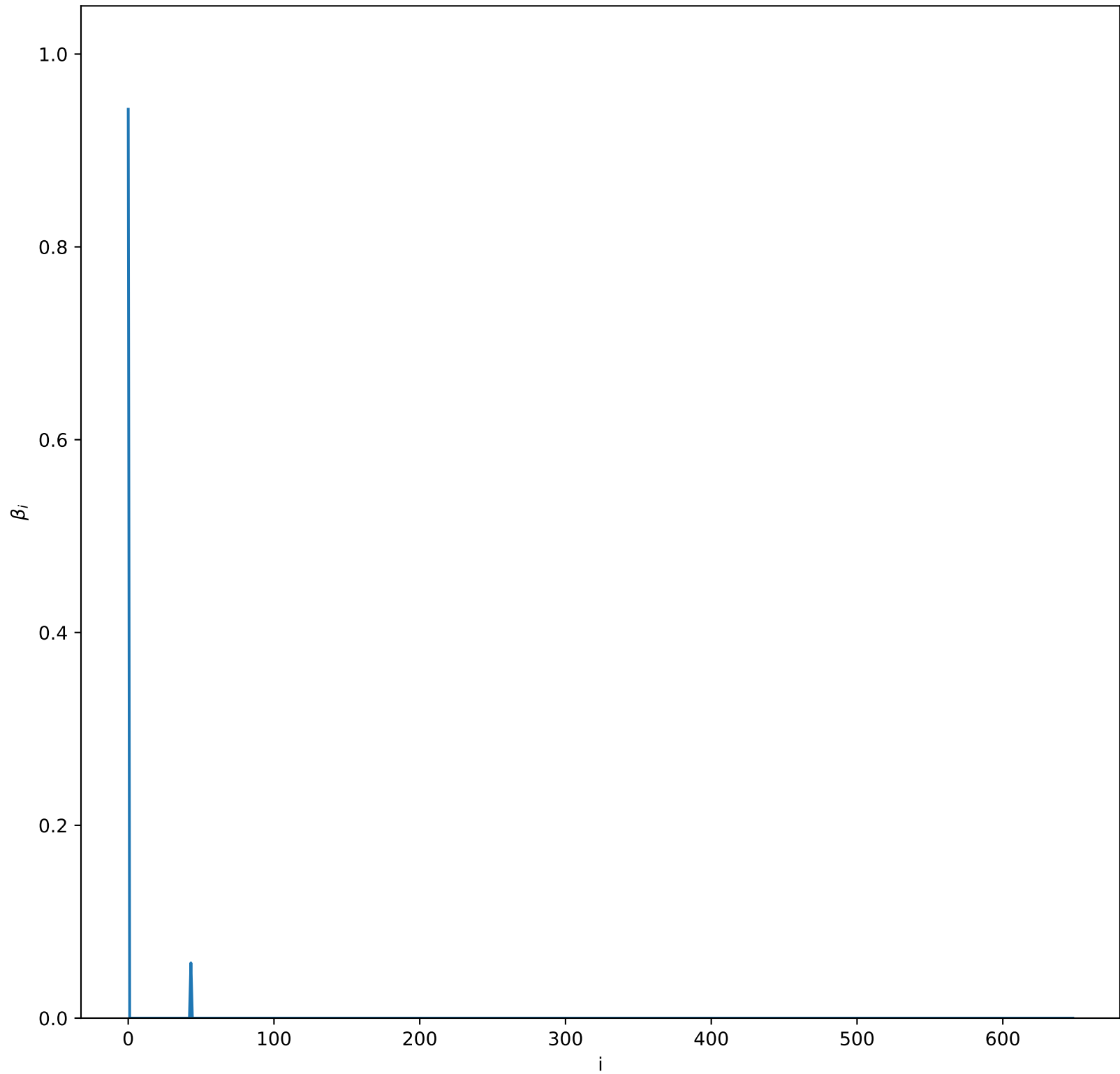
$\mu = 2.69$



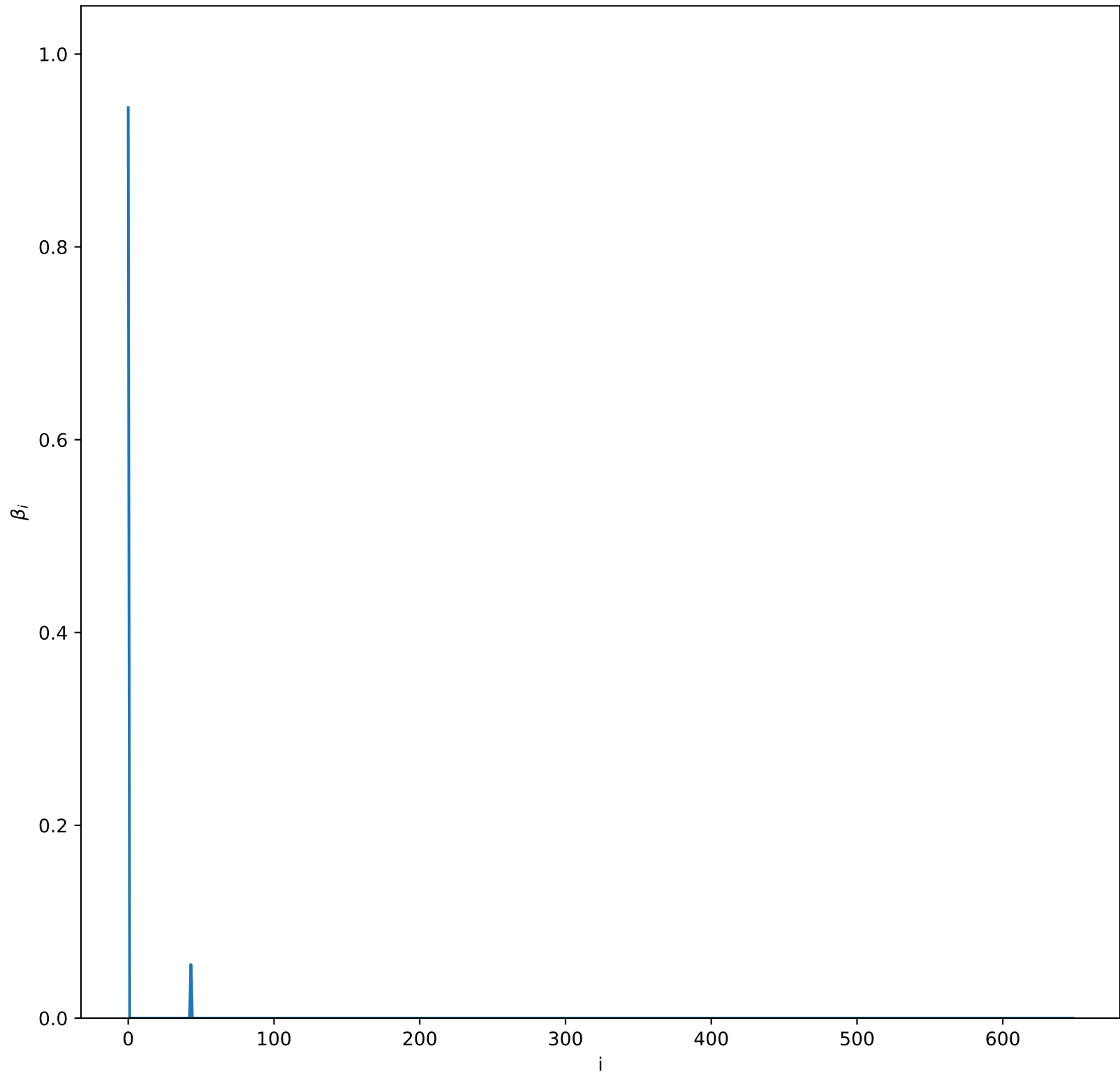
$\mu = 2.70$



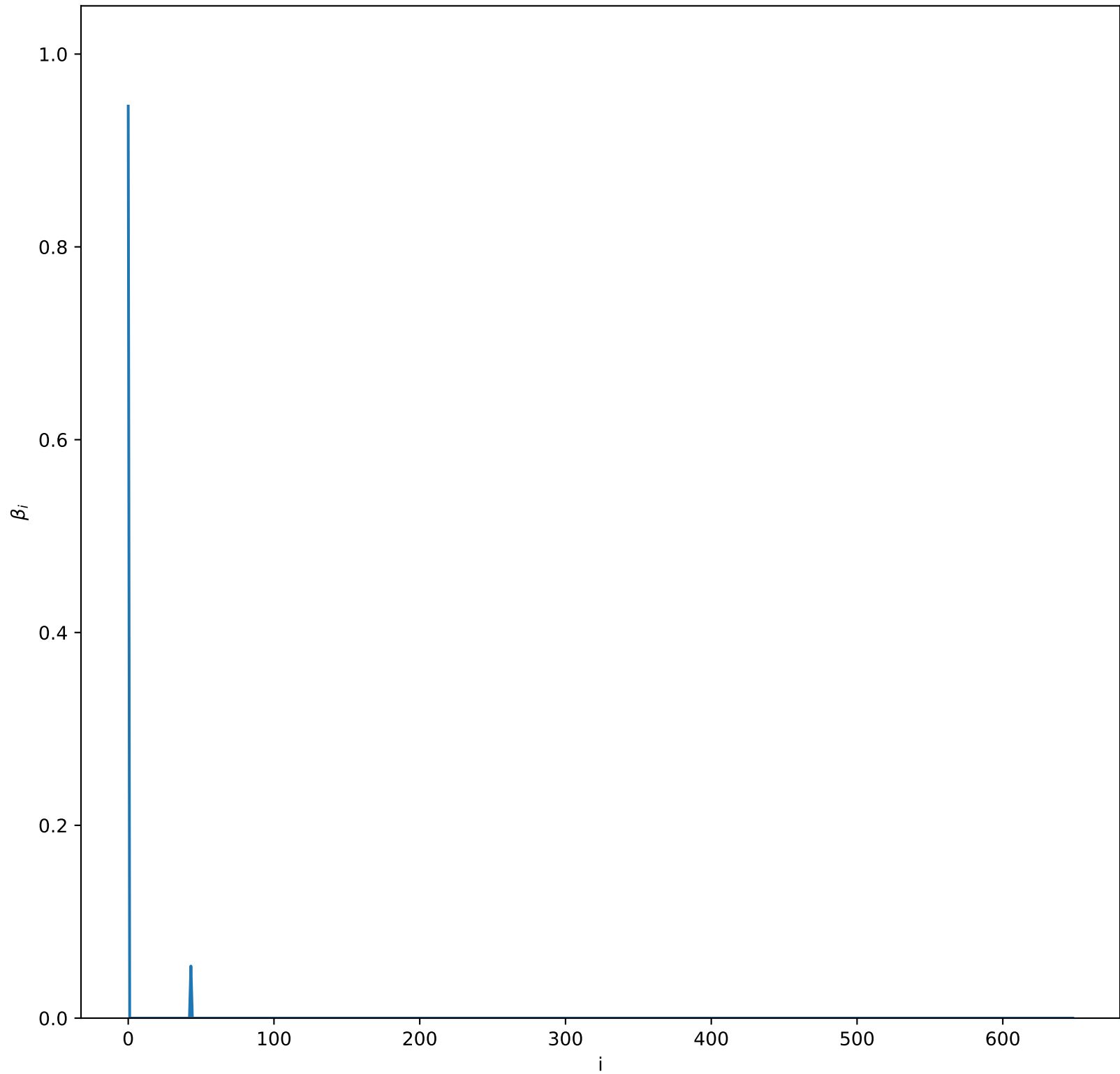
$\mu = 2.71$



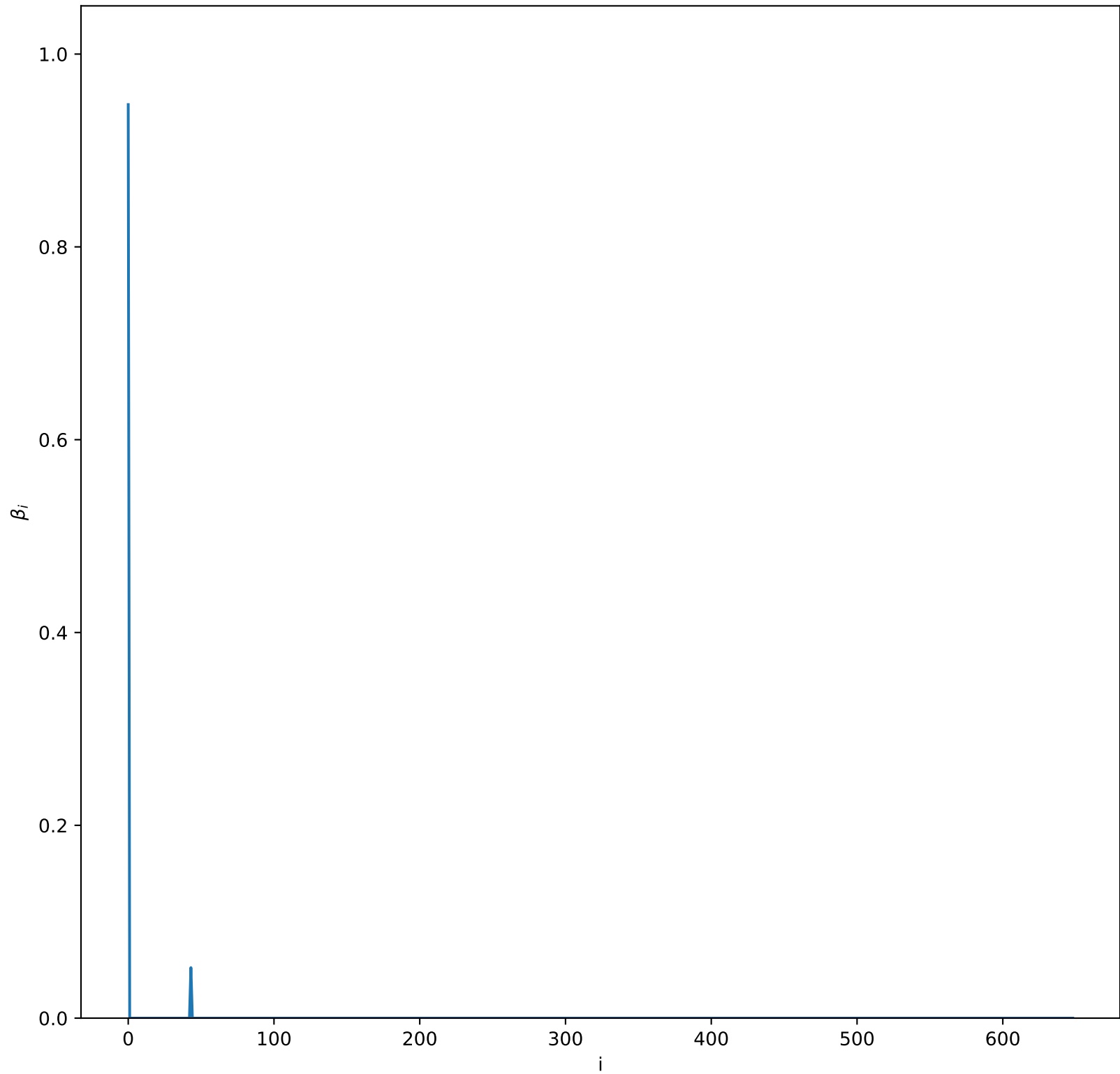
$\mu = 2.72$



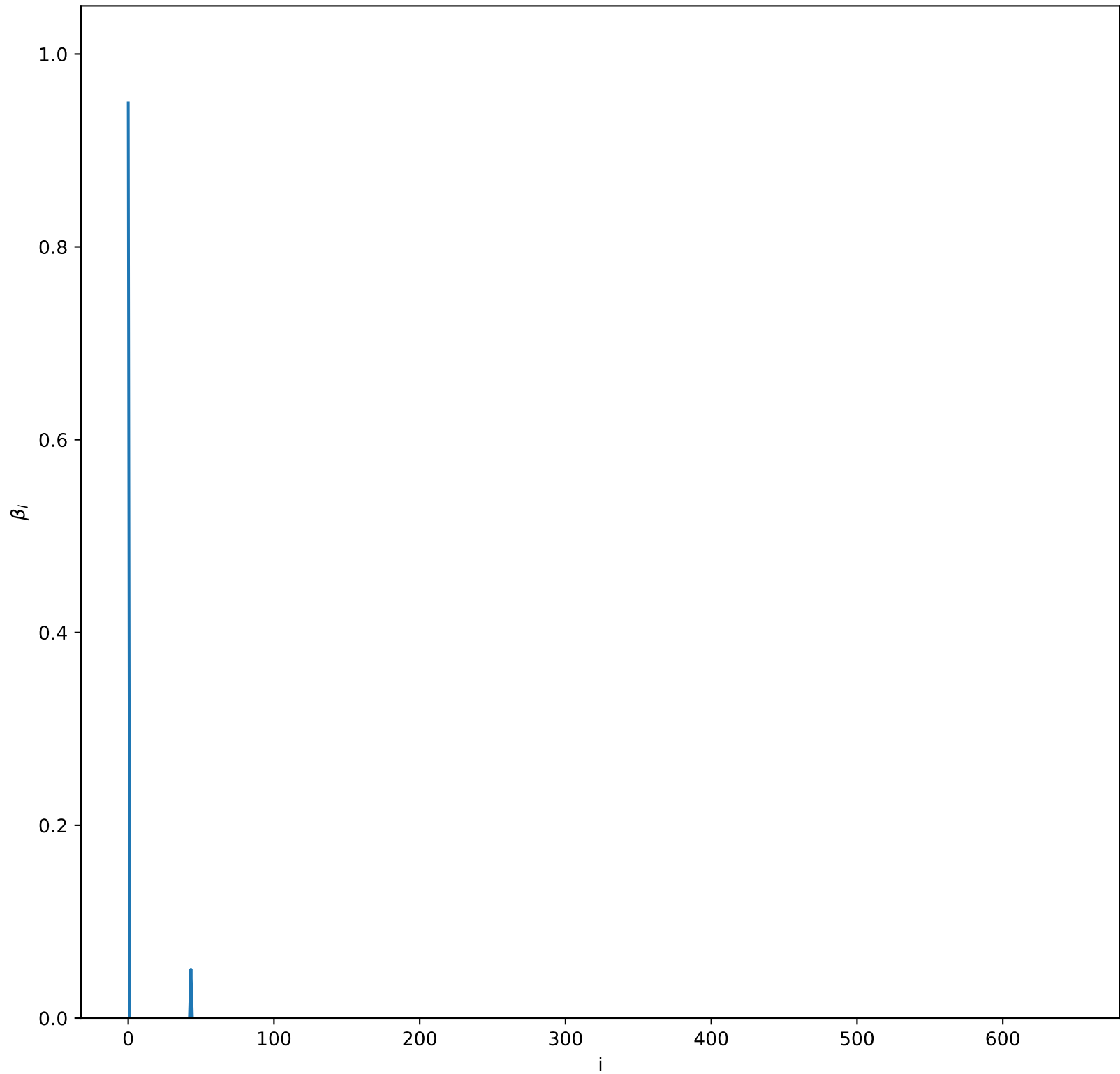
$\mu = 2.73$



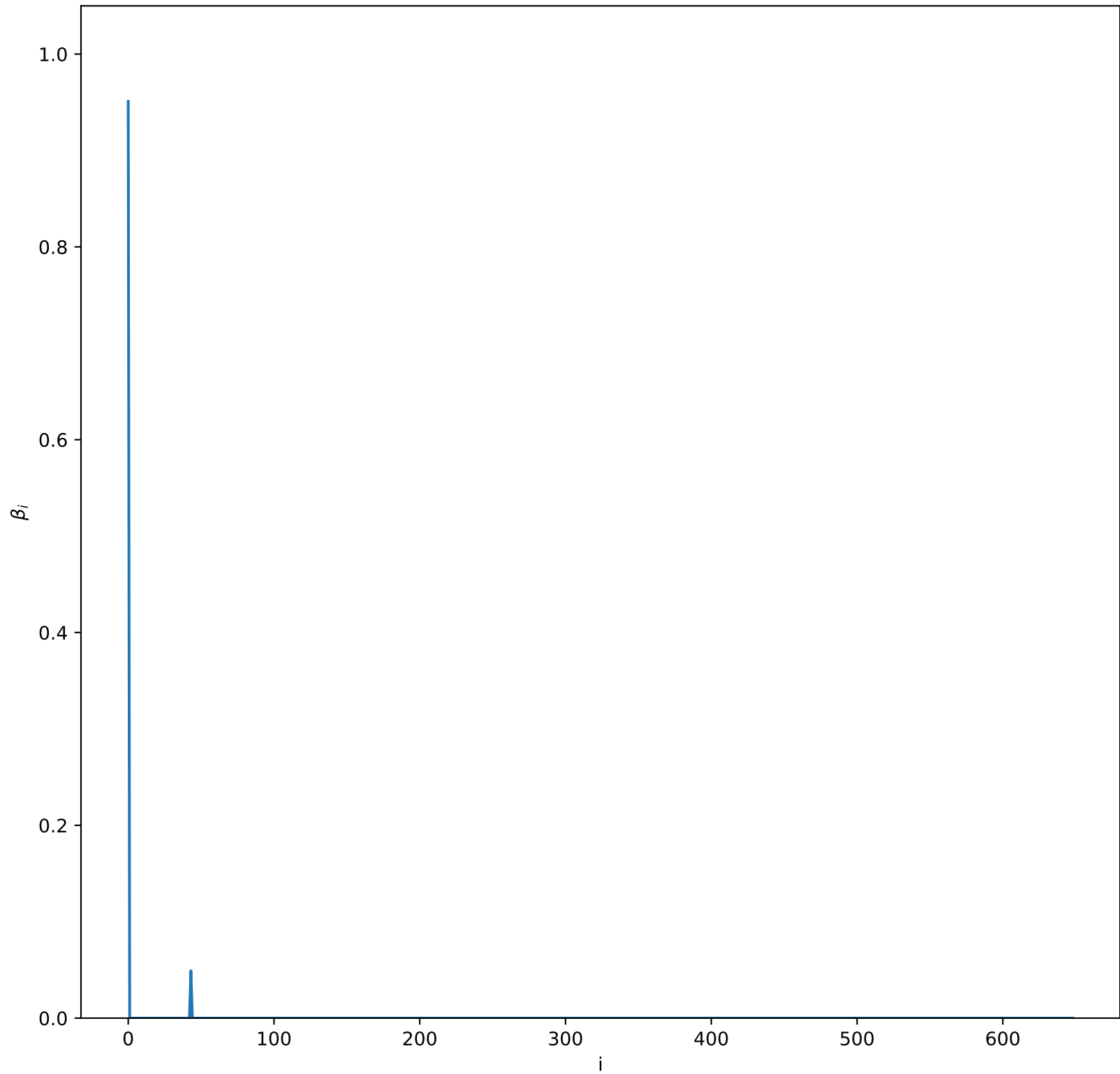
$\mu = 2.74$



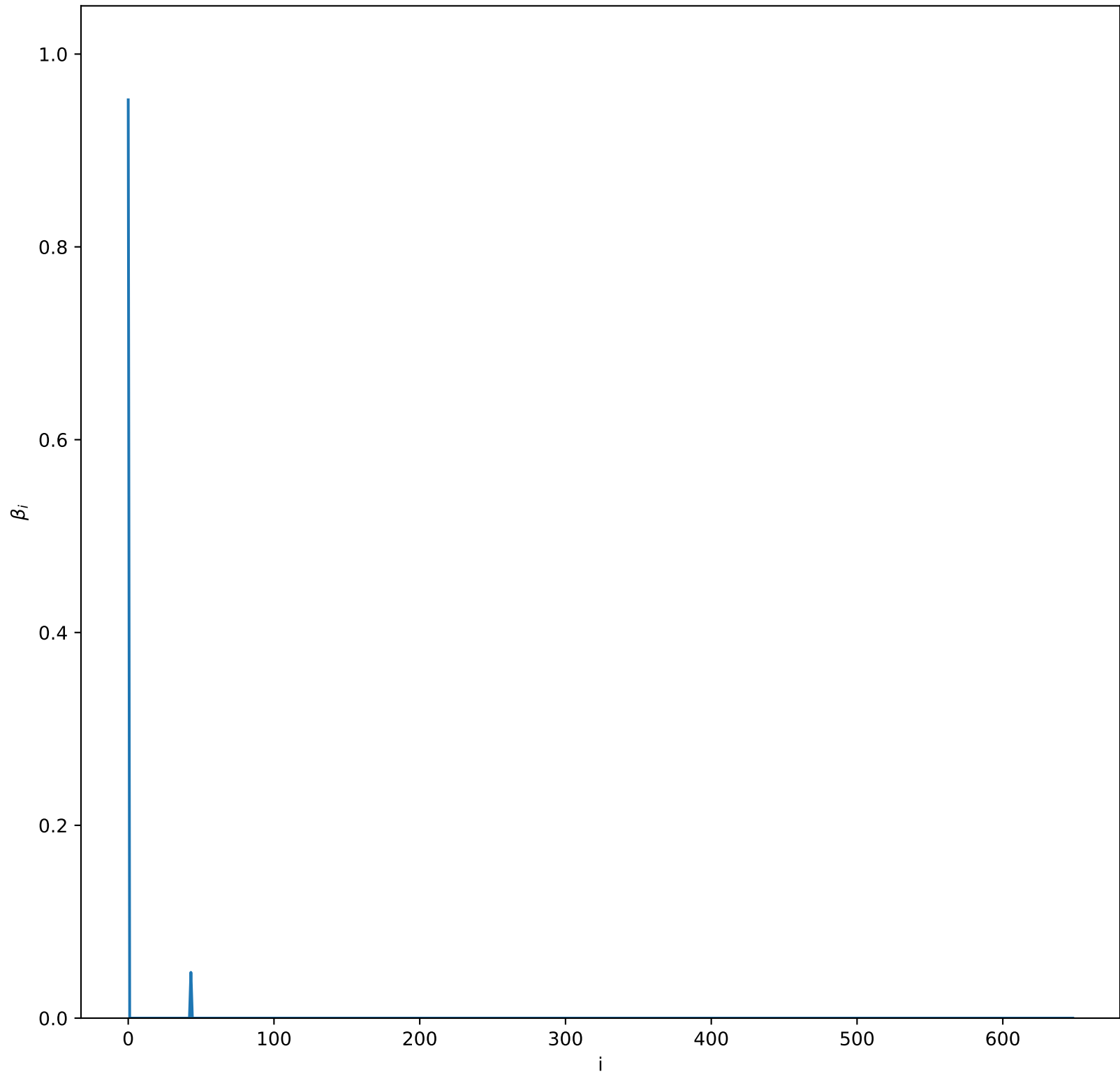
$\mu = 2.75$



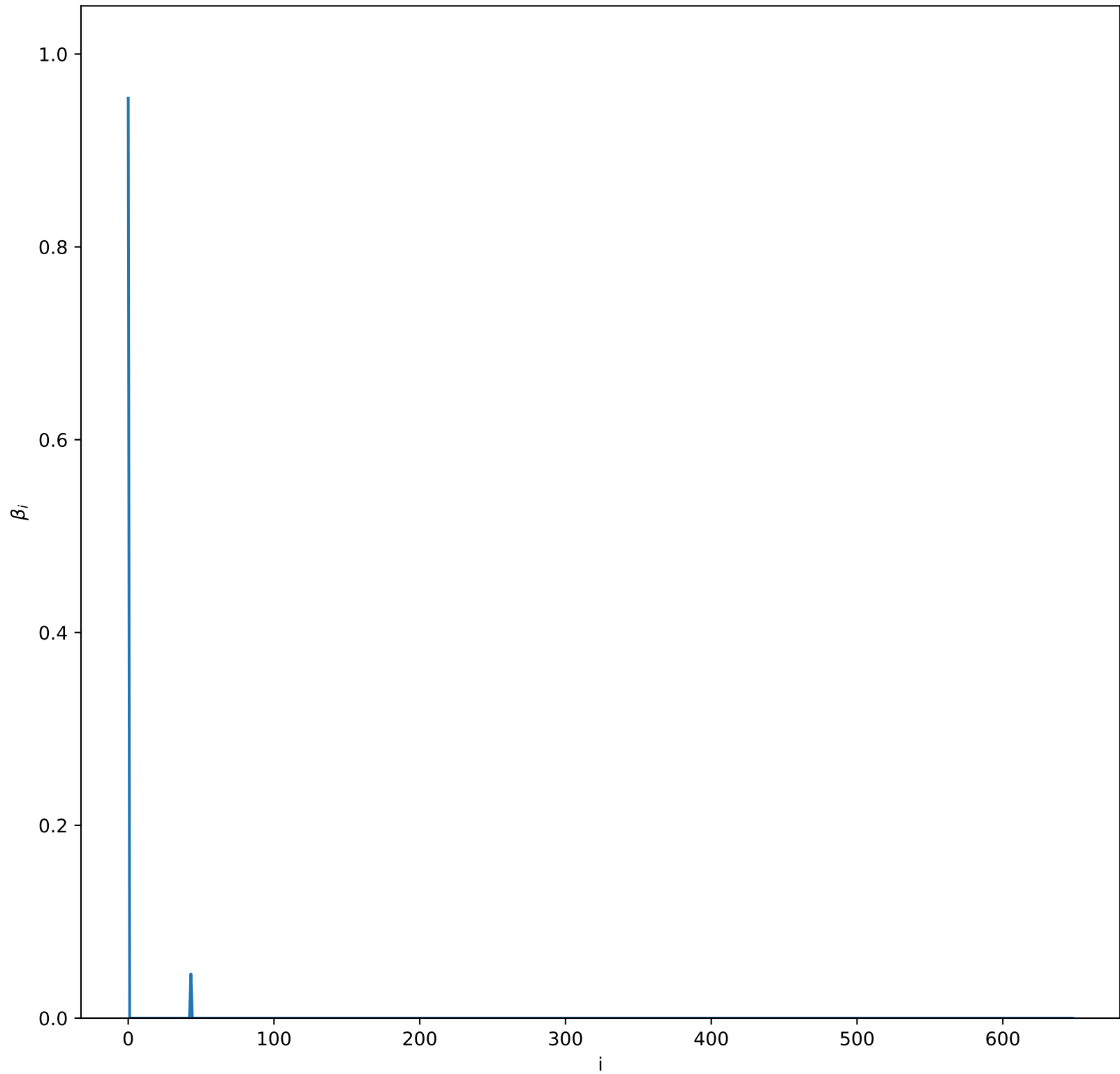
$\mu = 2.76$



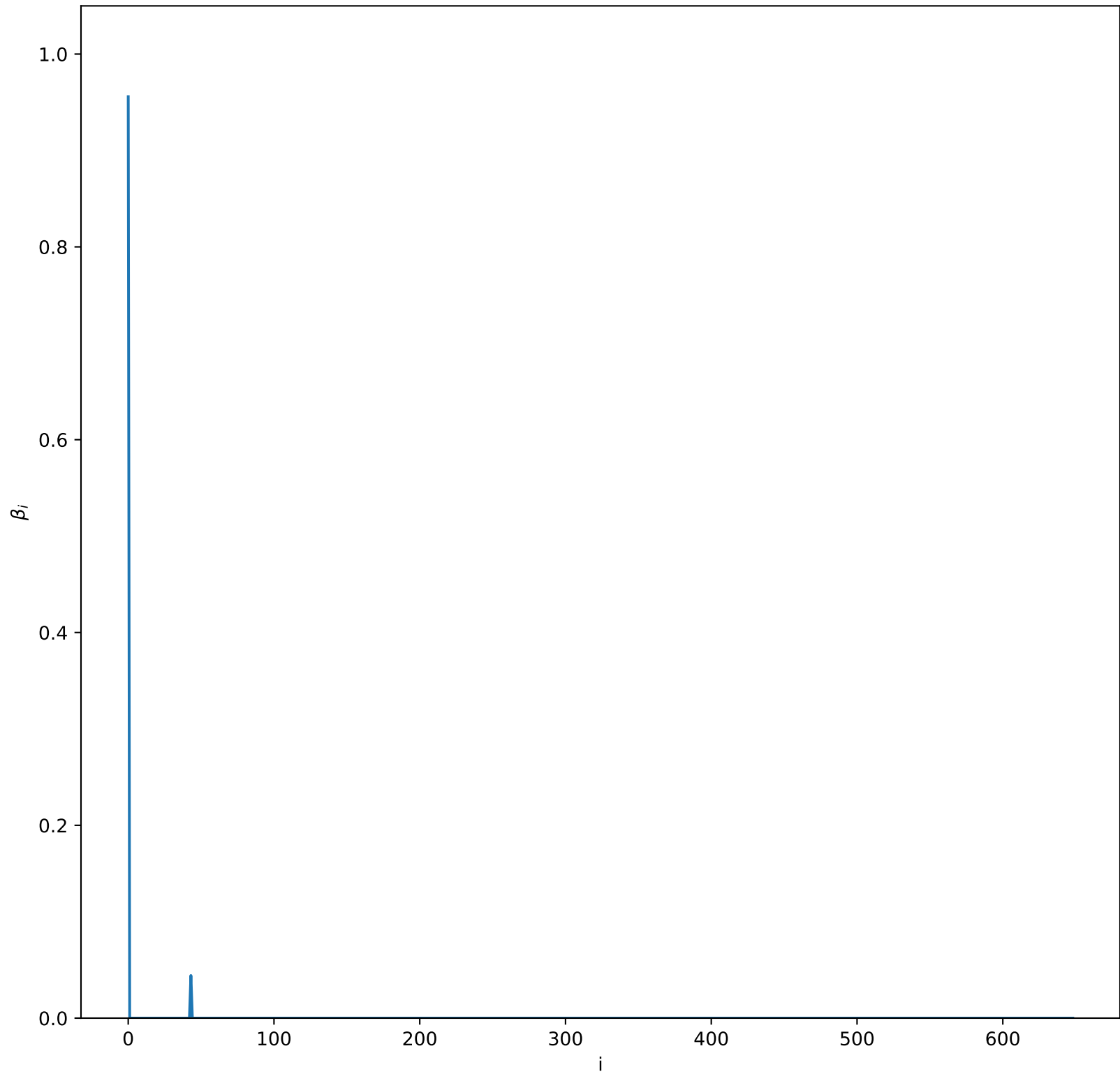
$\mu = 2.77$



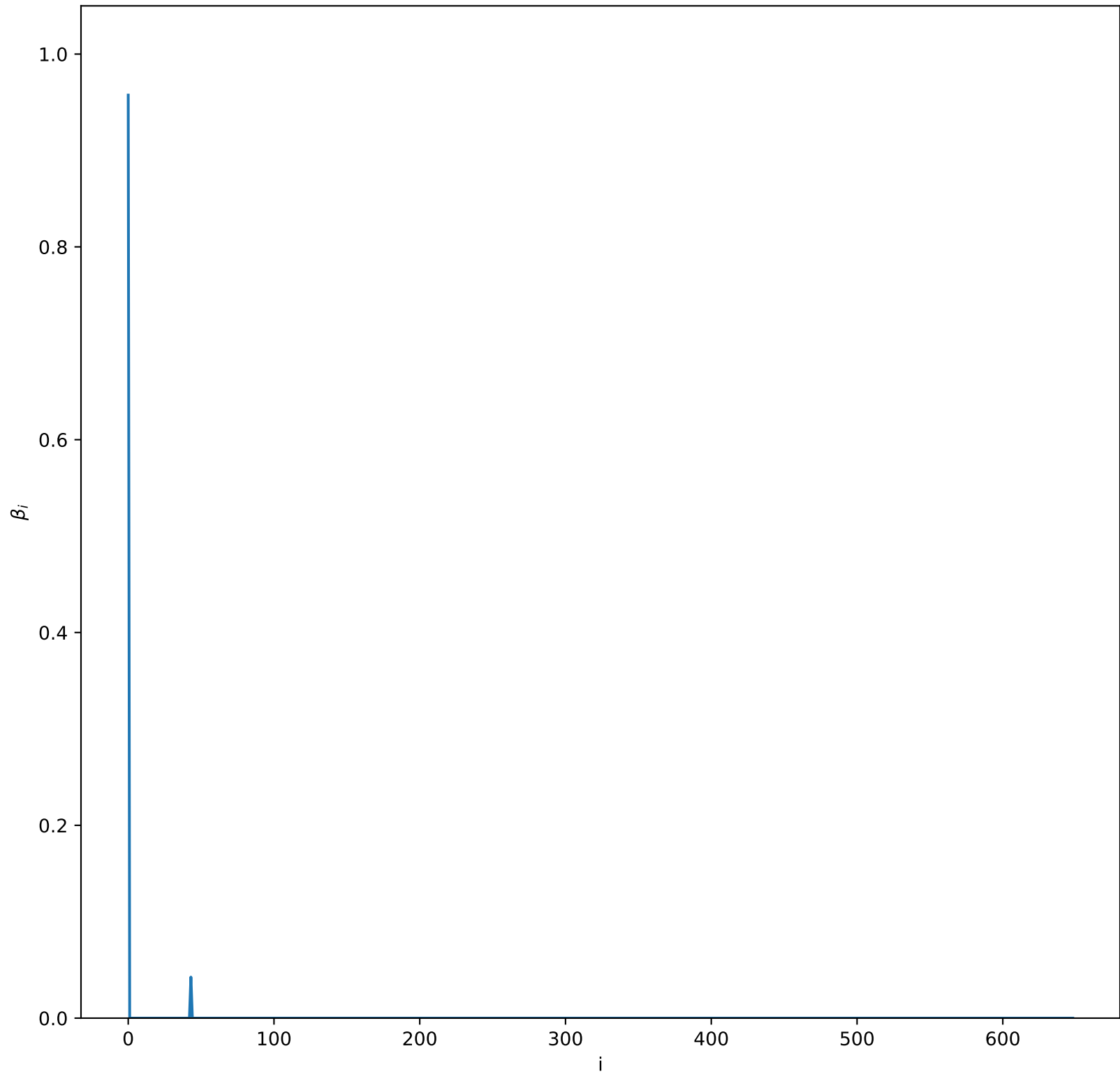
$\mu = 2.78$



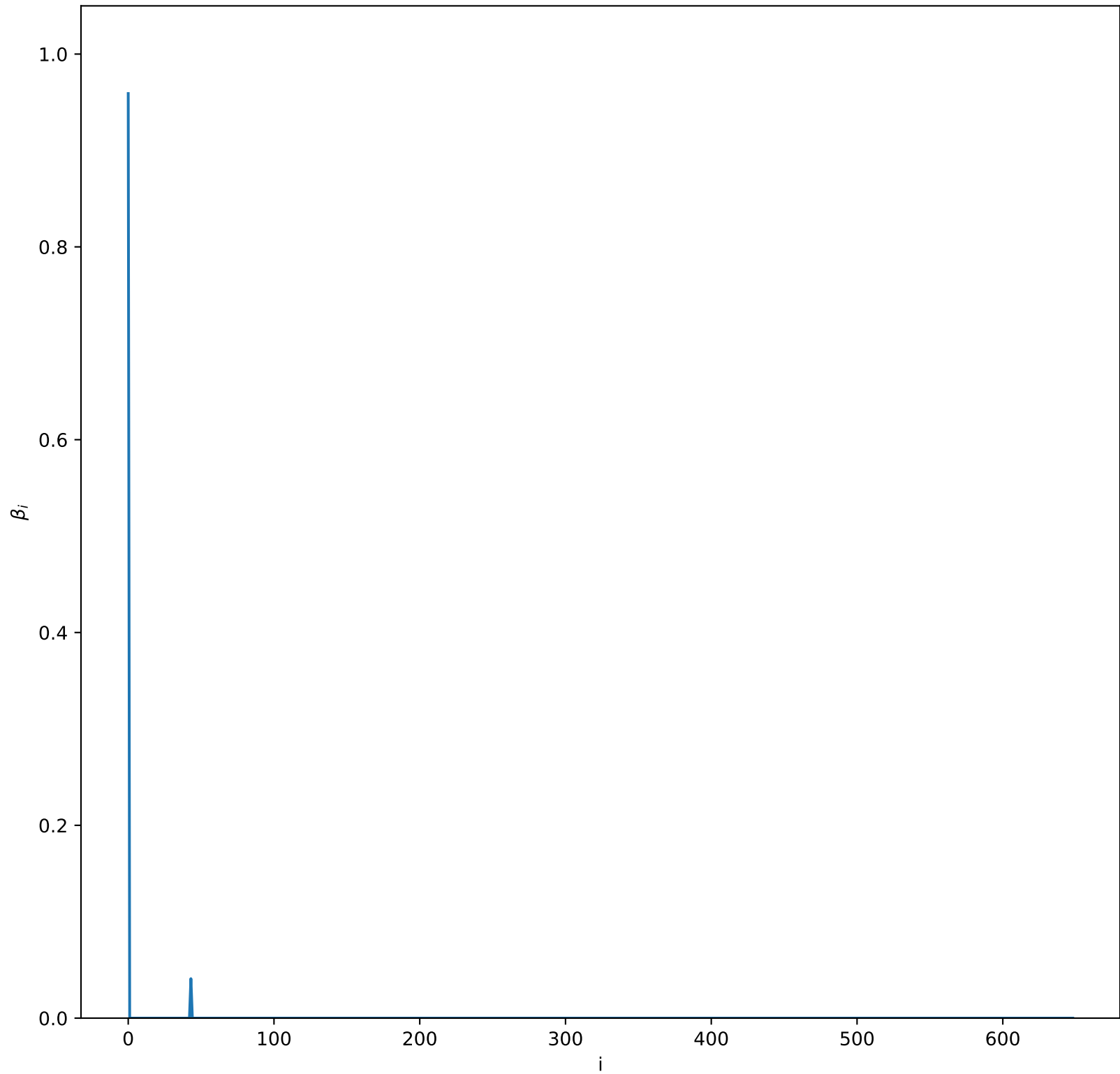
$\mu = 2.79$



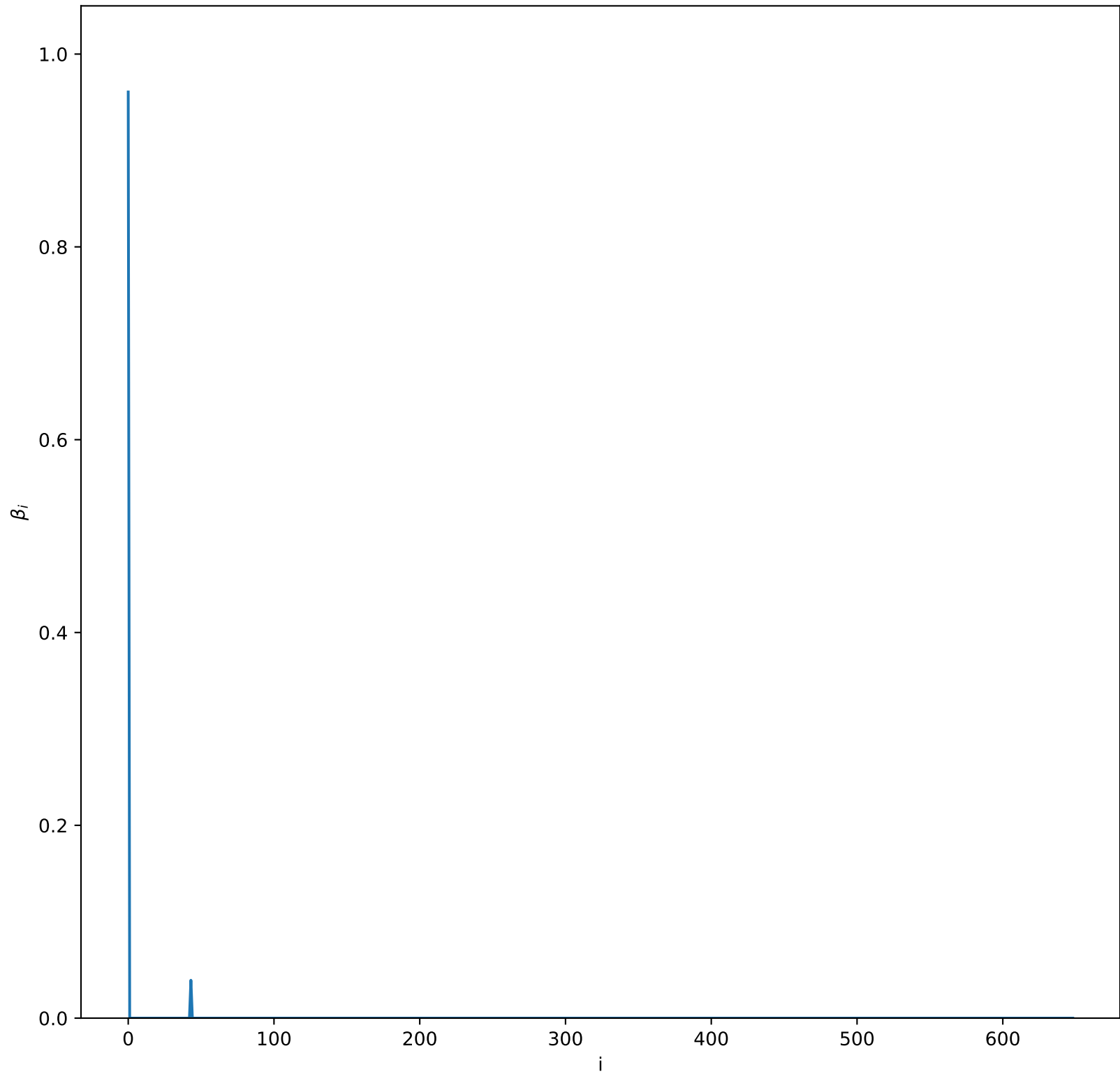
$\mu = 2.80$



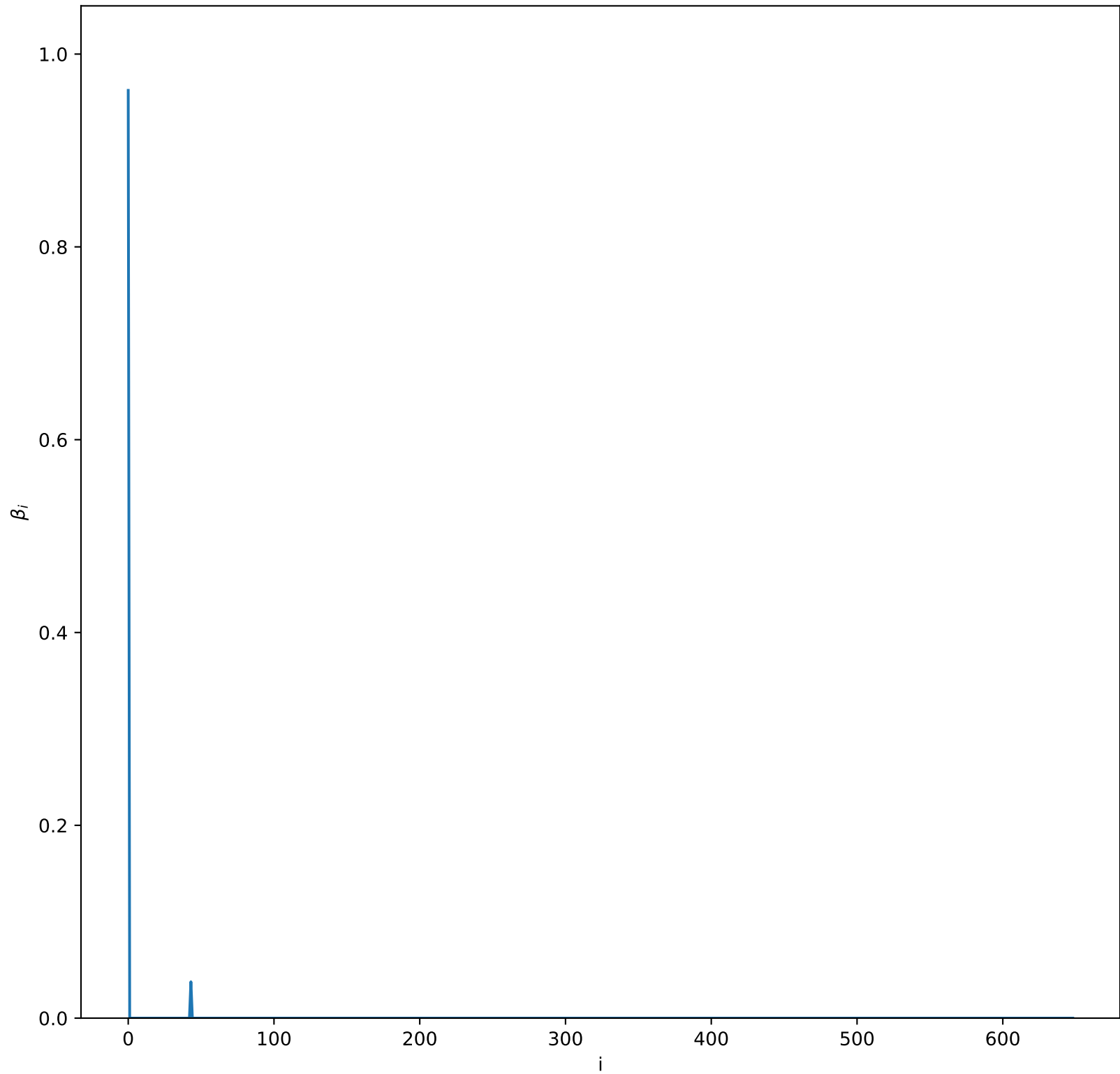
$\mu = 2.81$



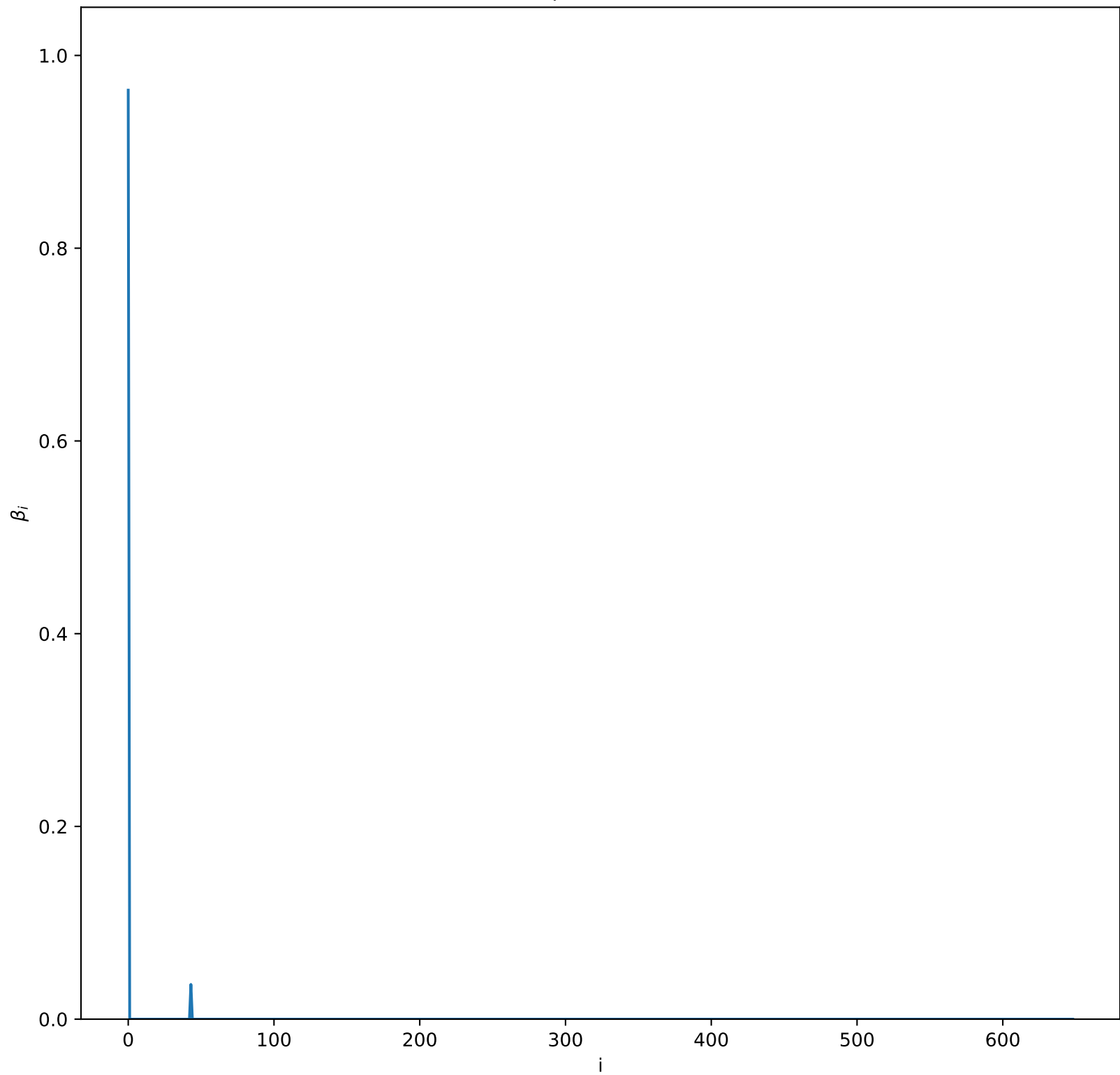
$\mu = 2.82$



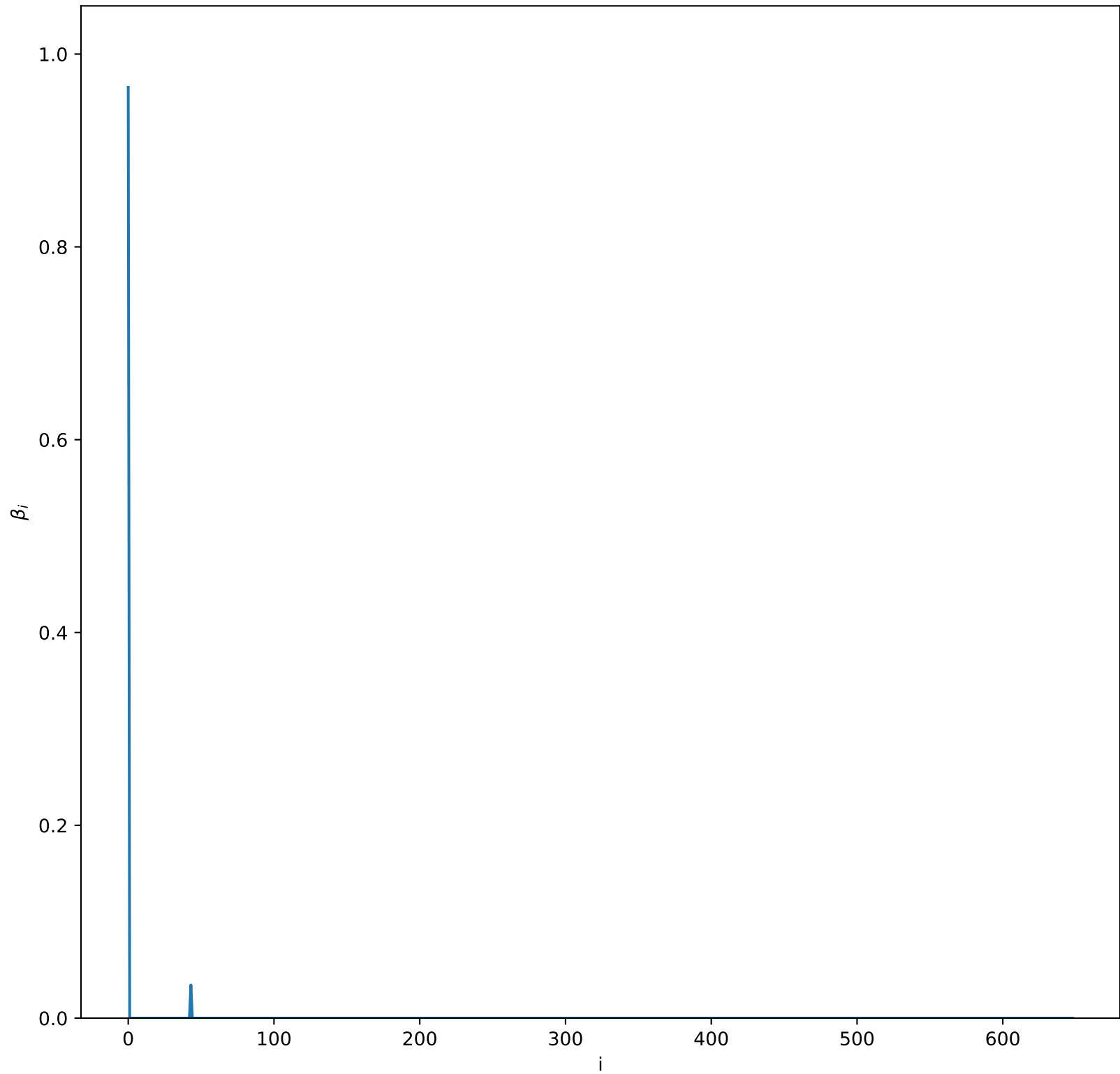
$\mu = 2.83$



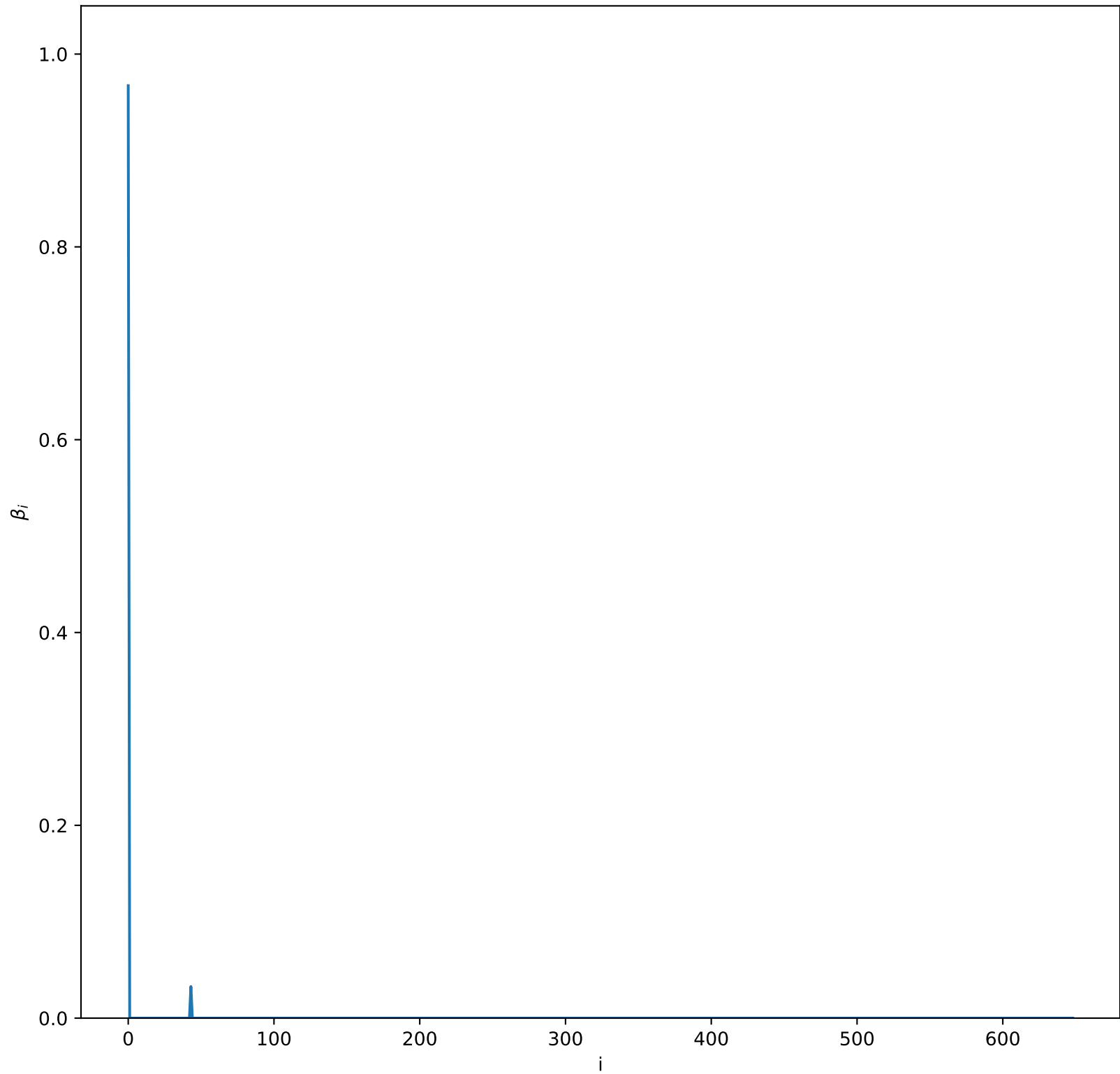
$\mu = 2.84$



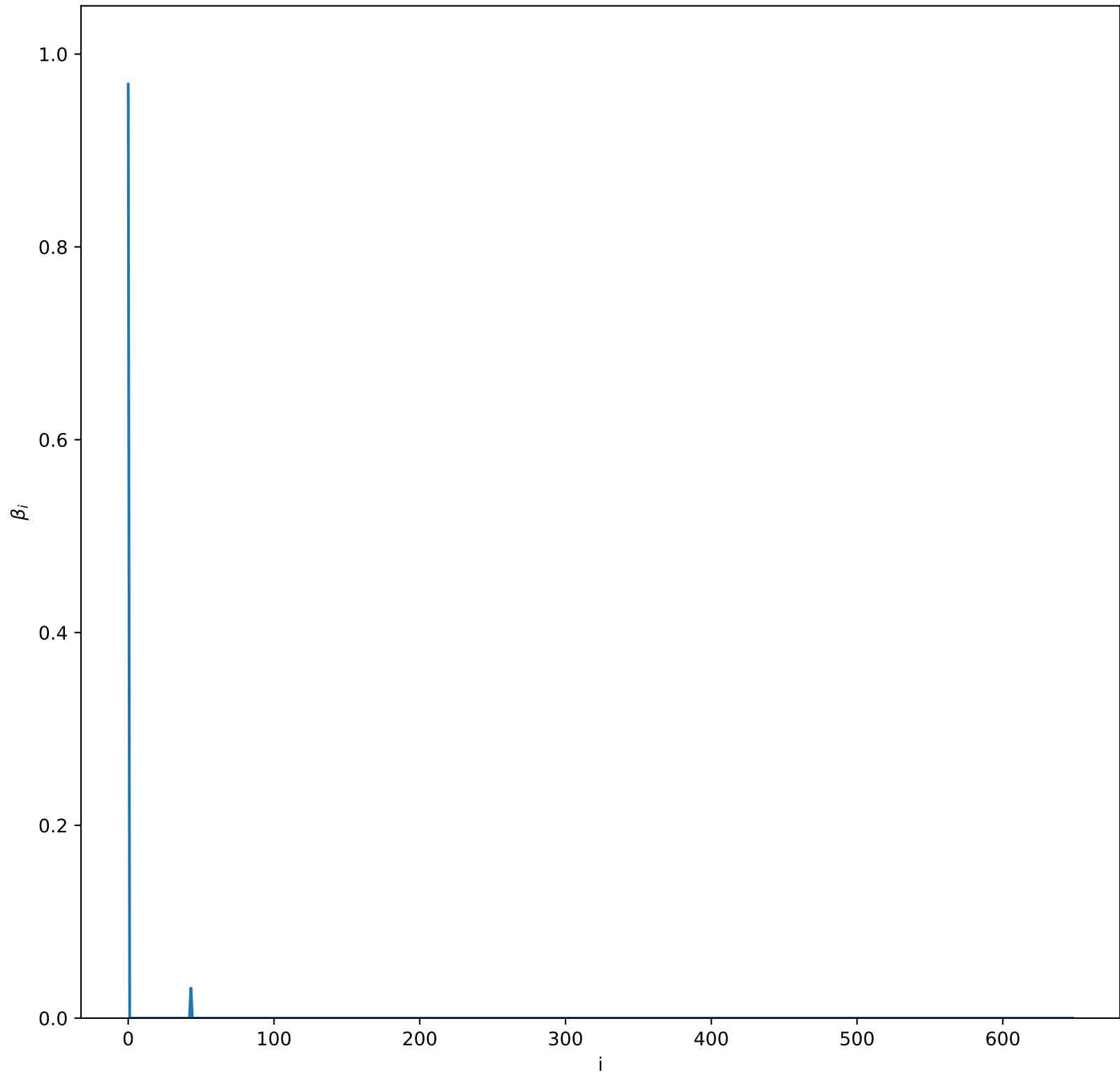
$\mu = 2.85$



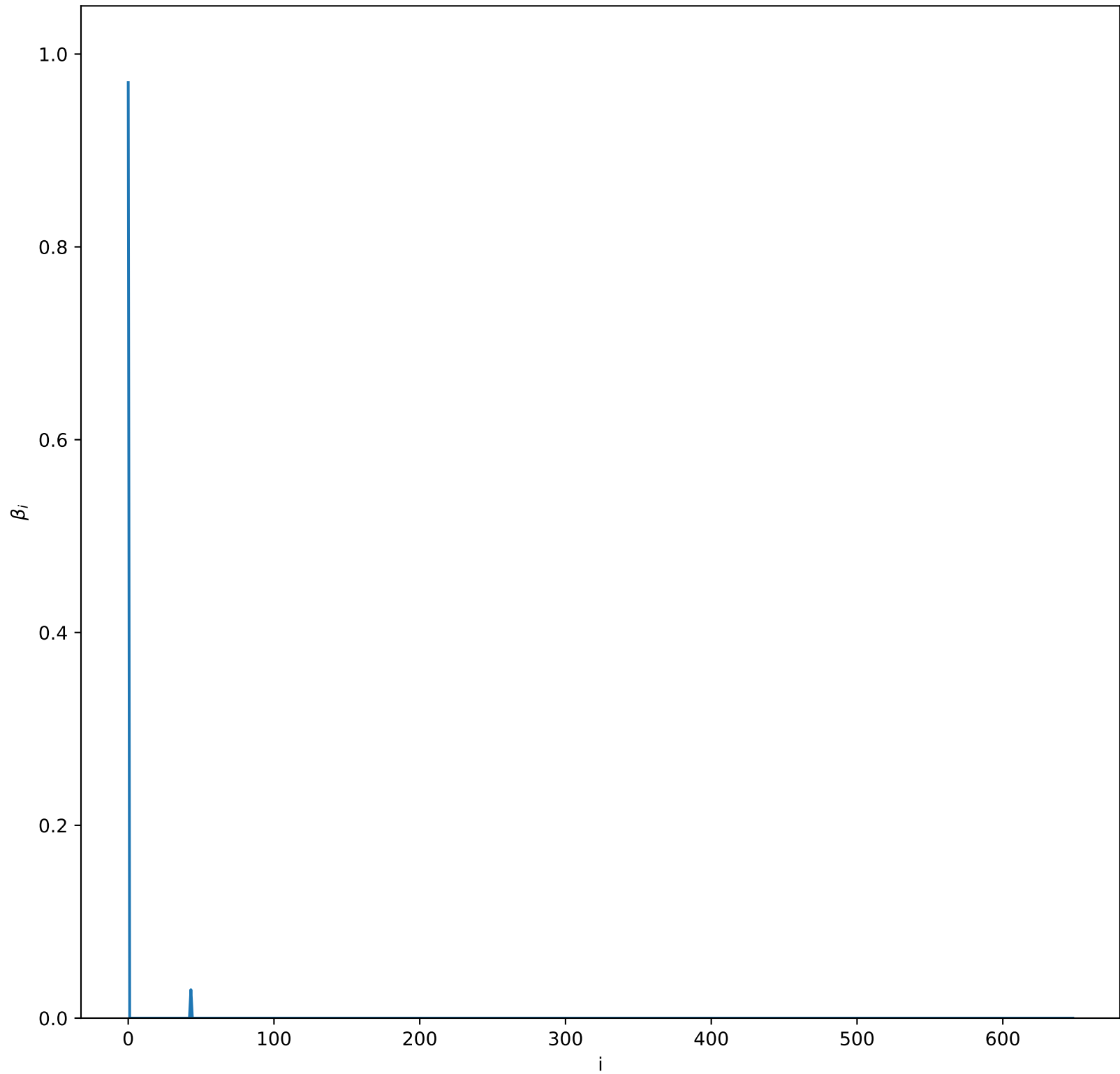
$\mu = 2.86$



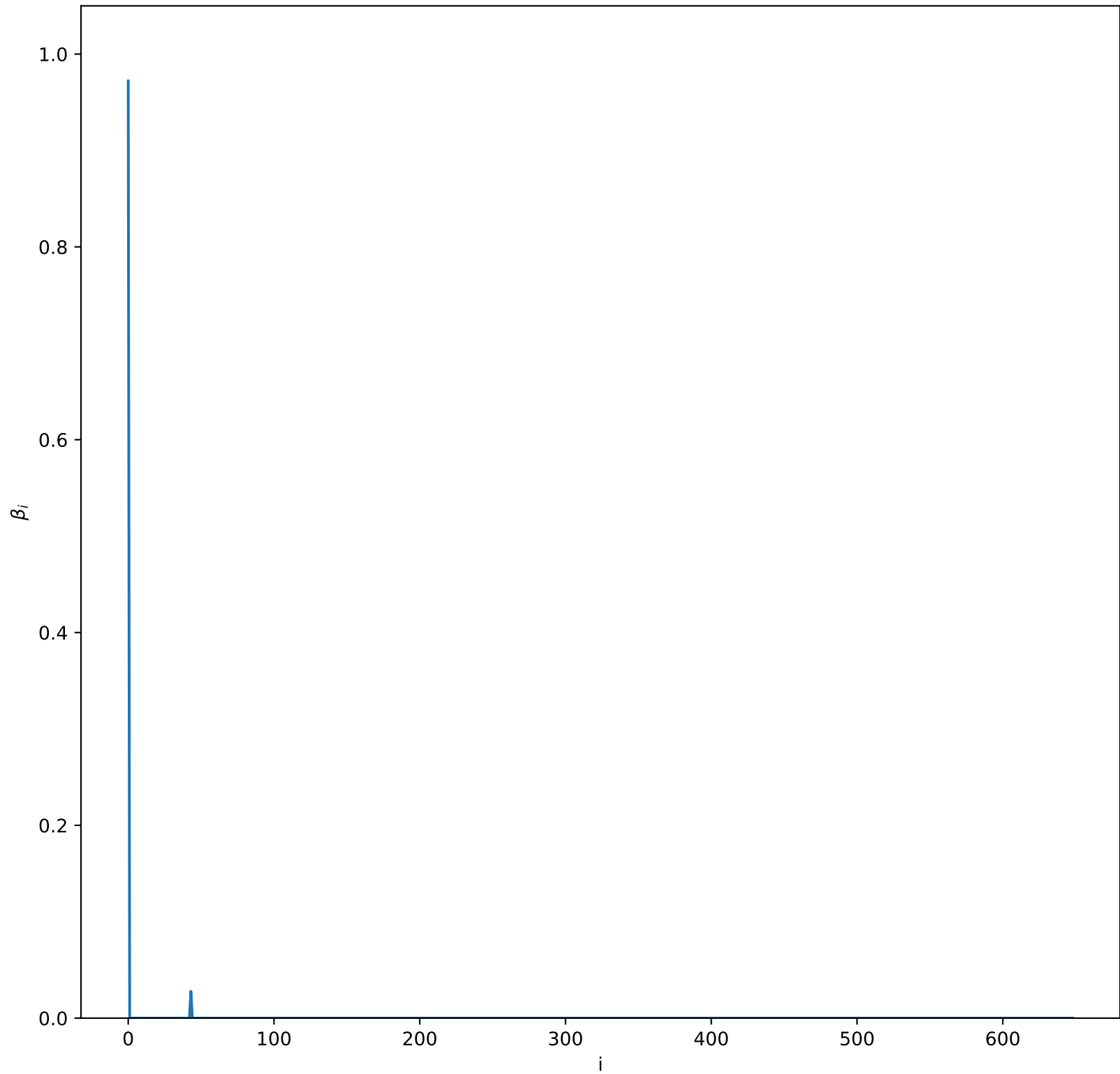
$\mu = 2.87$



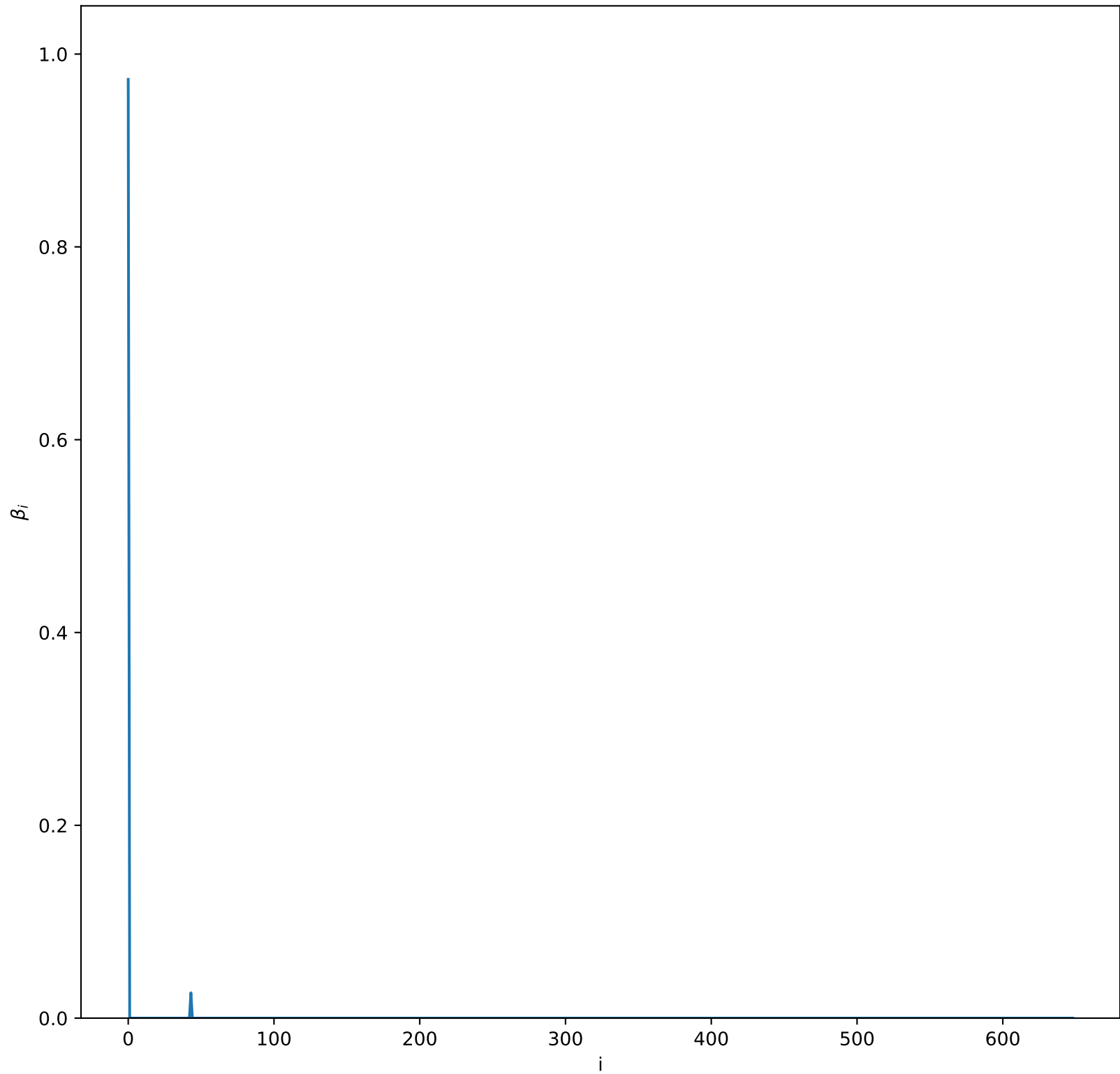
$\mu = 2.88$



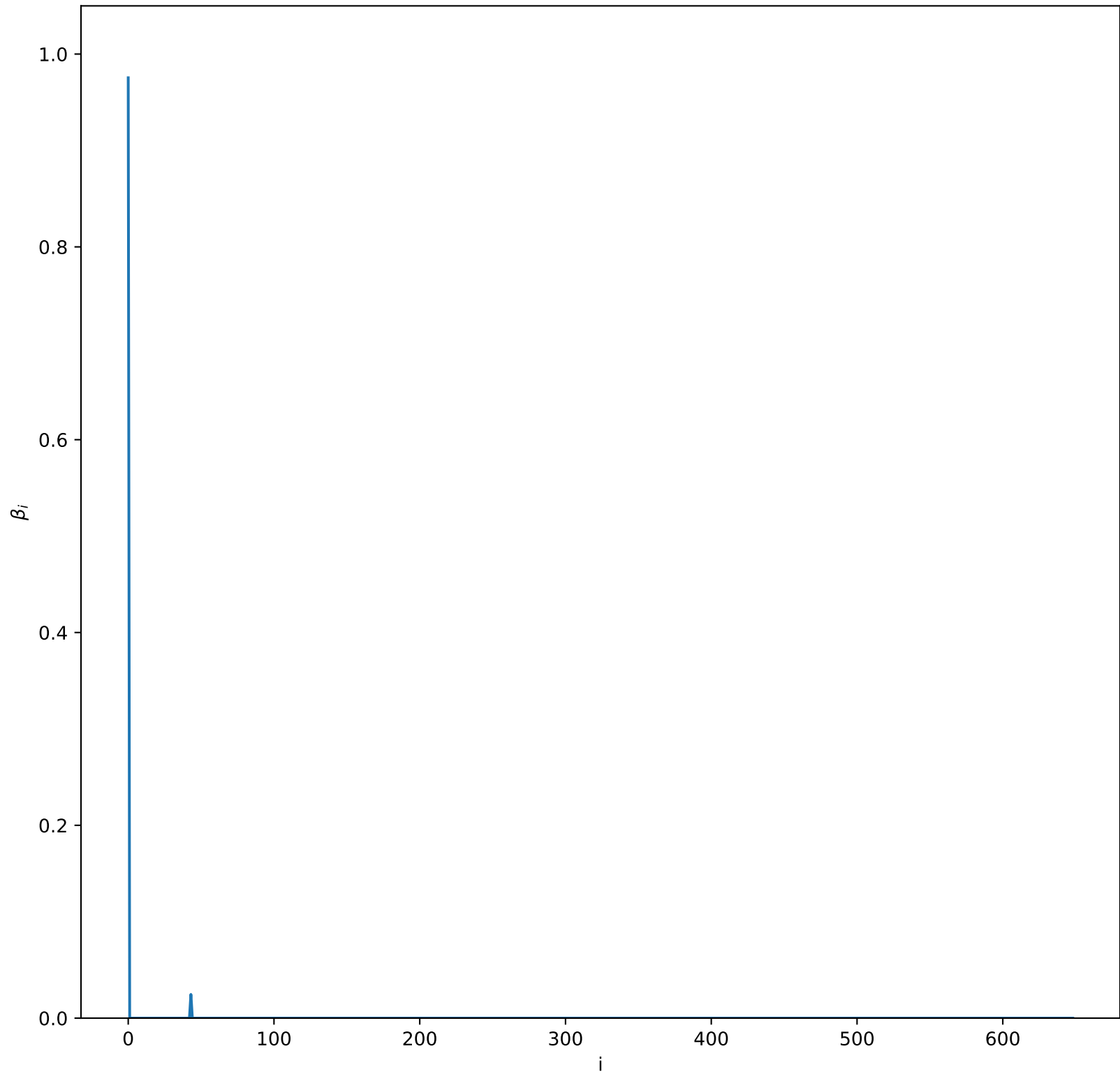
$\mu = 2.89$



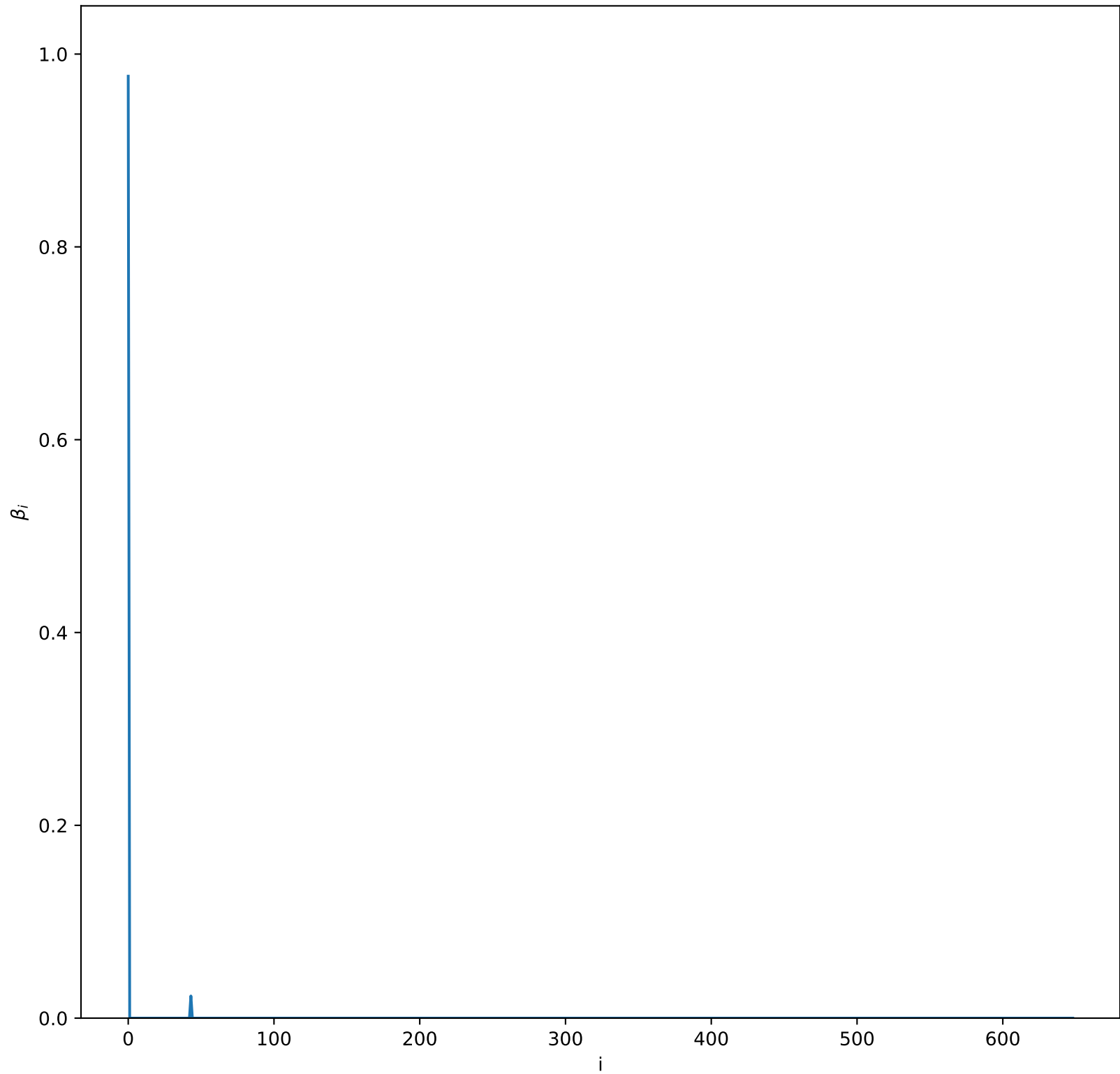
$\mu = 2.90$



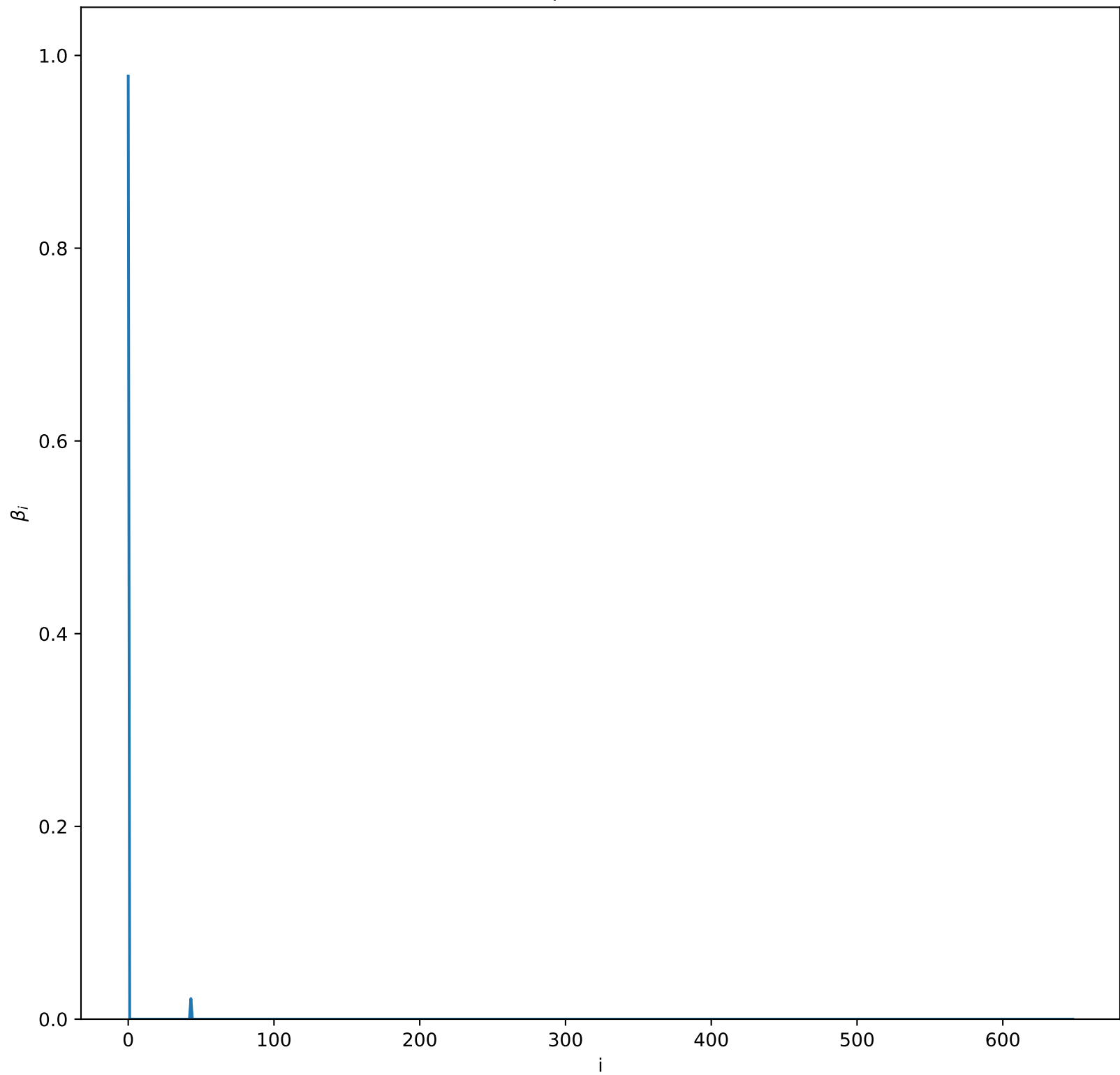
$\mu = 2.91$



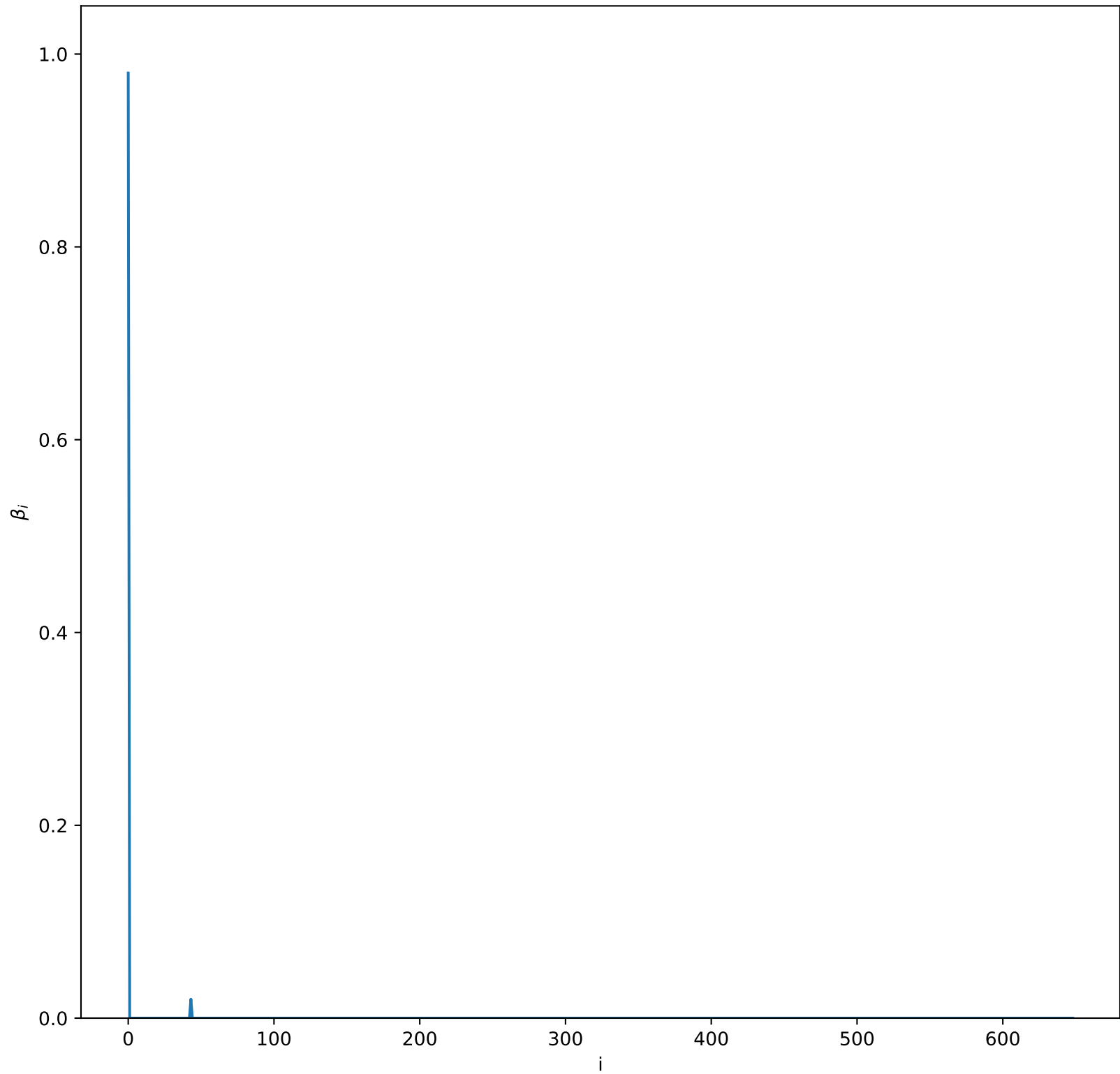
$\mu = 2.92$



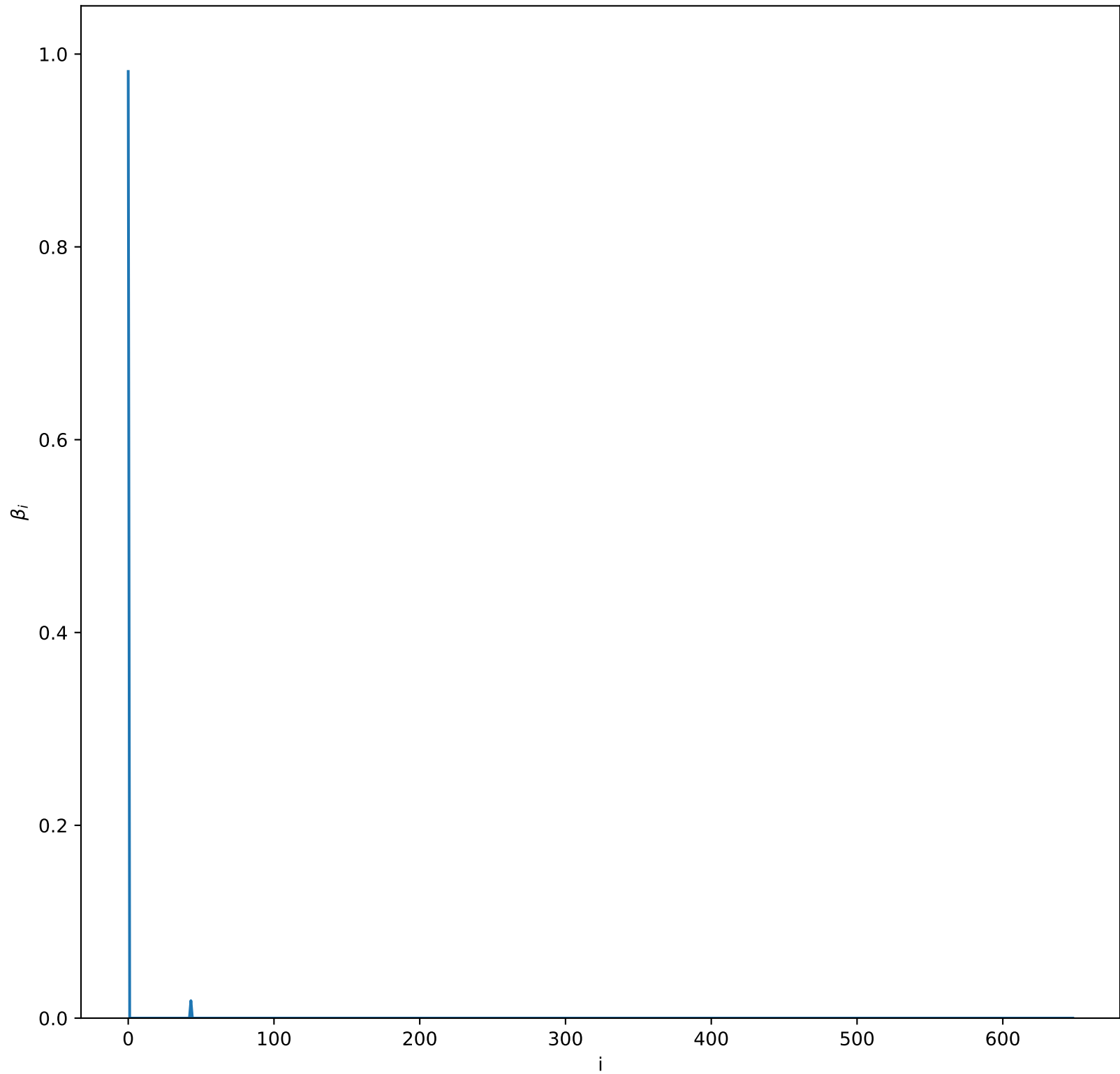
$\mu = 2.93$



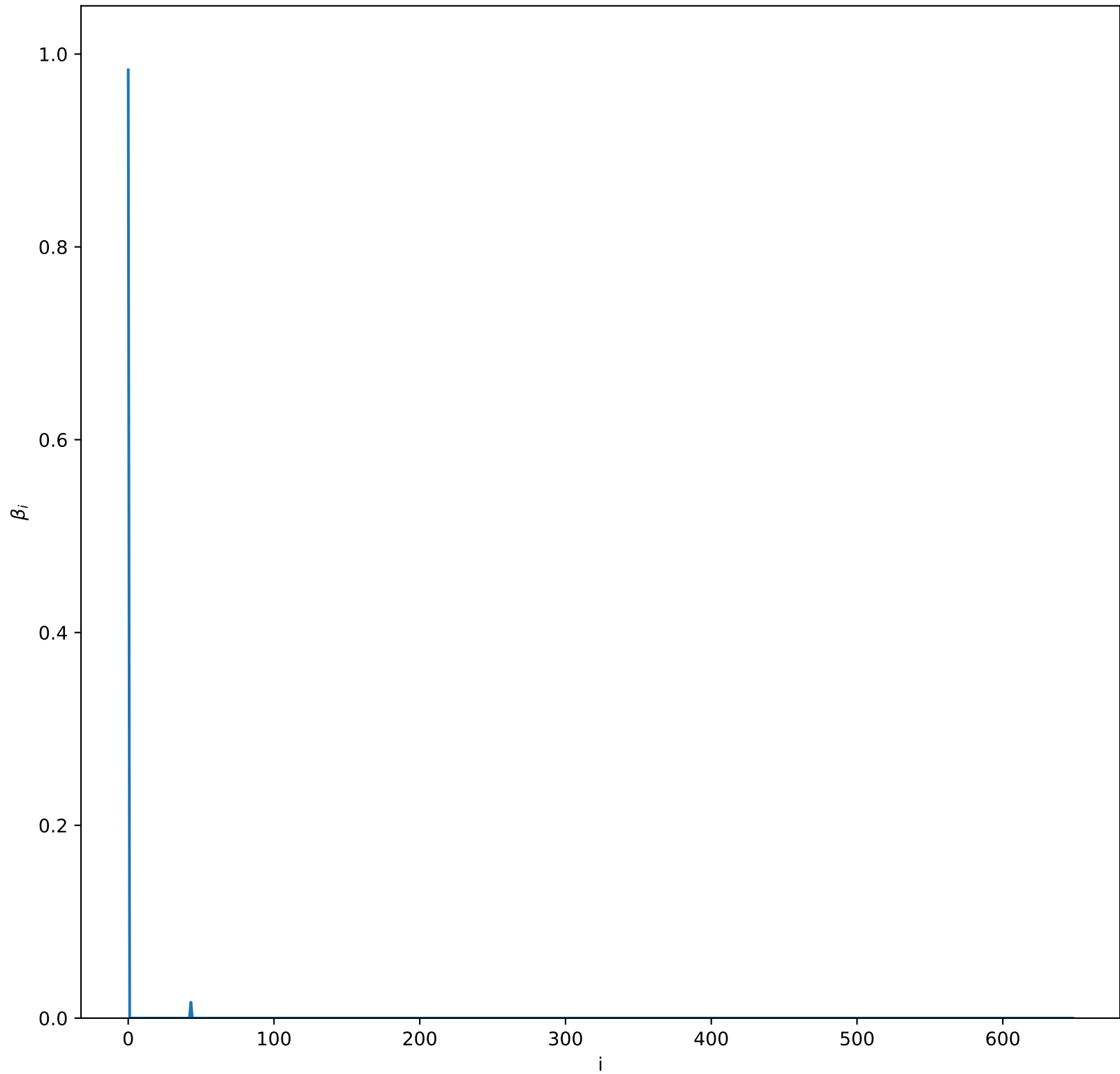
$\mu = 2.94$



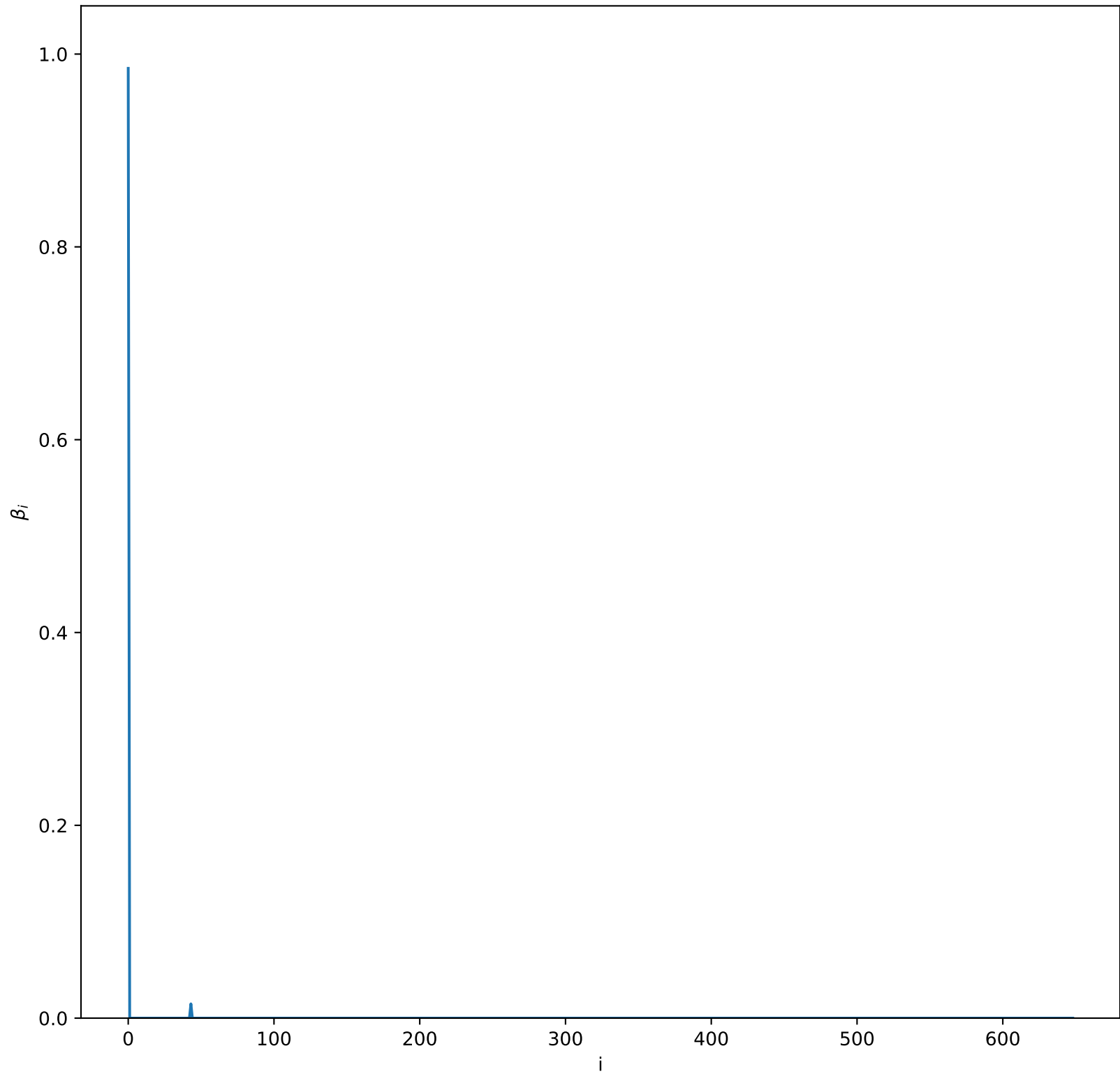
$\mu = 2.95$



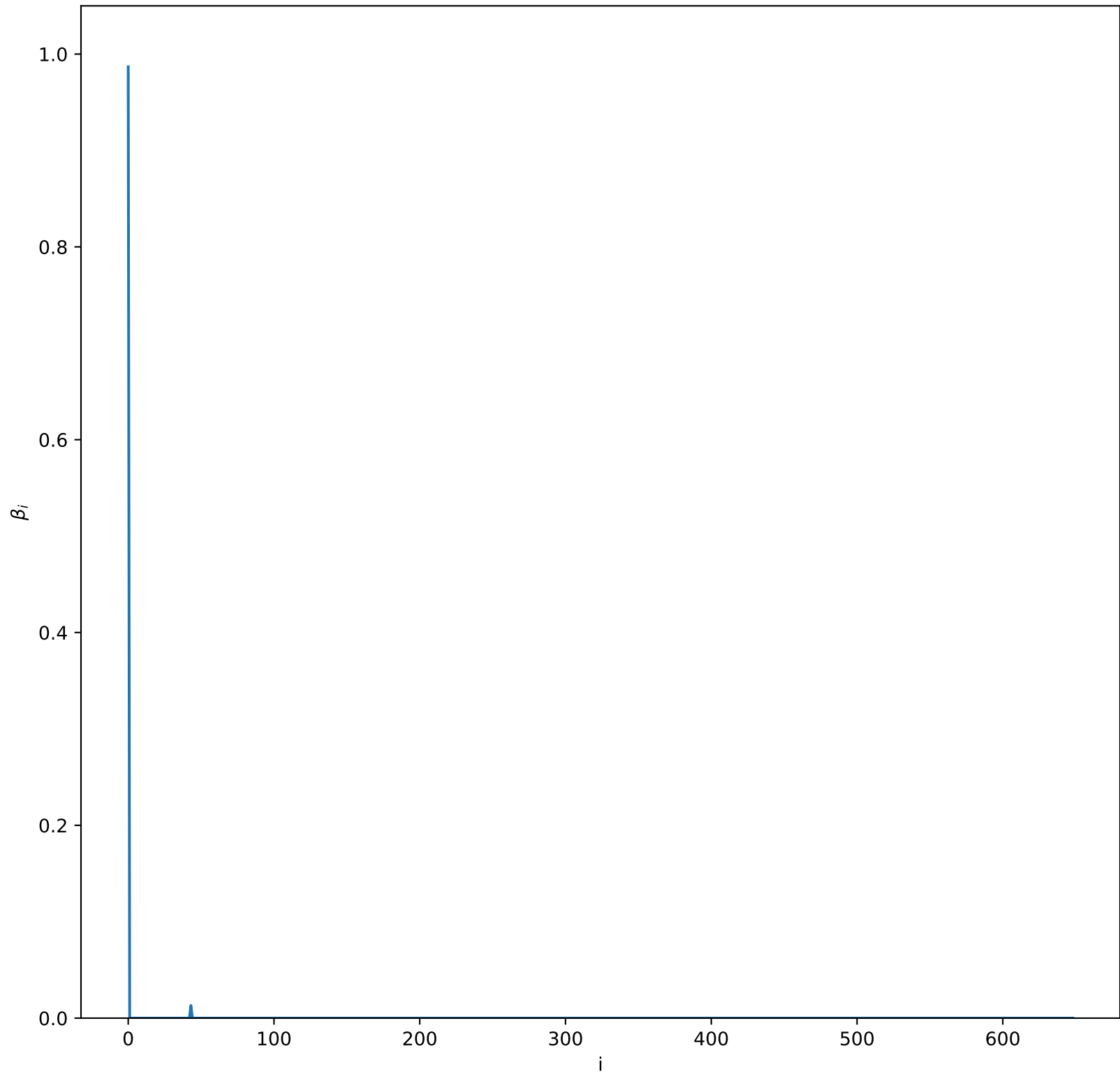
$\mu = 2.96$



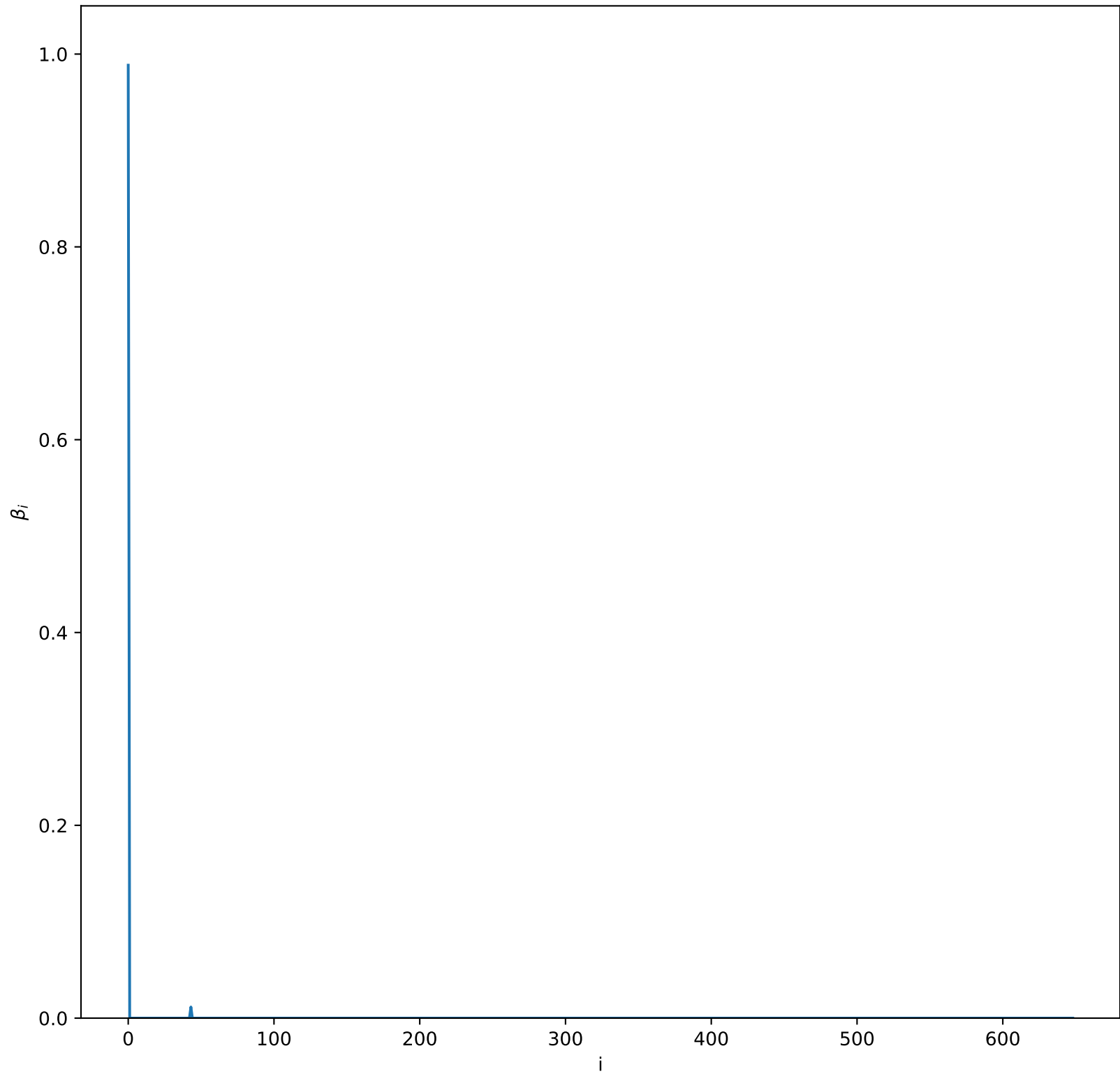
$\mu = 2.97$



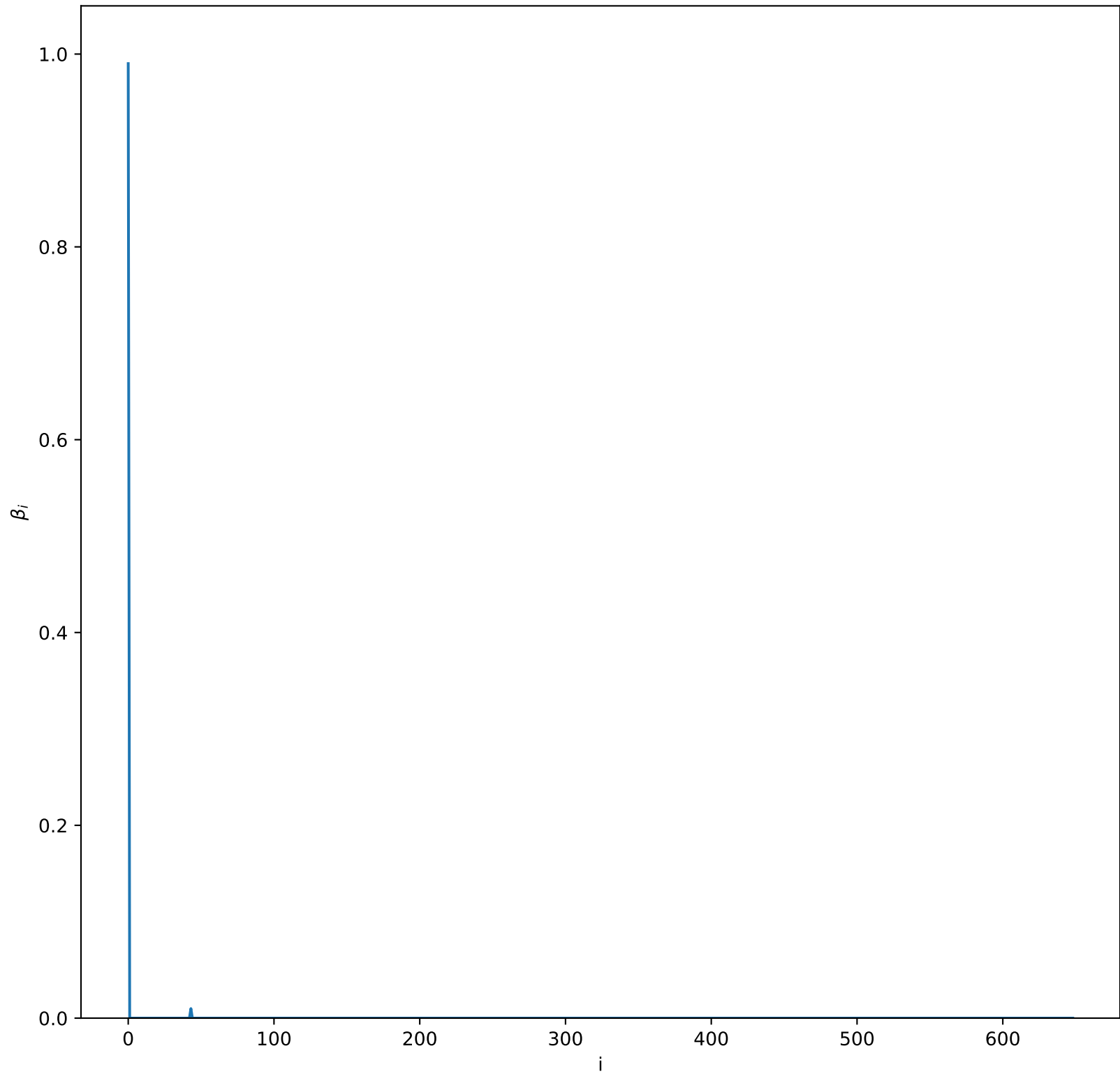
$\mu = 2.98$



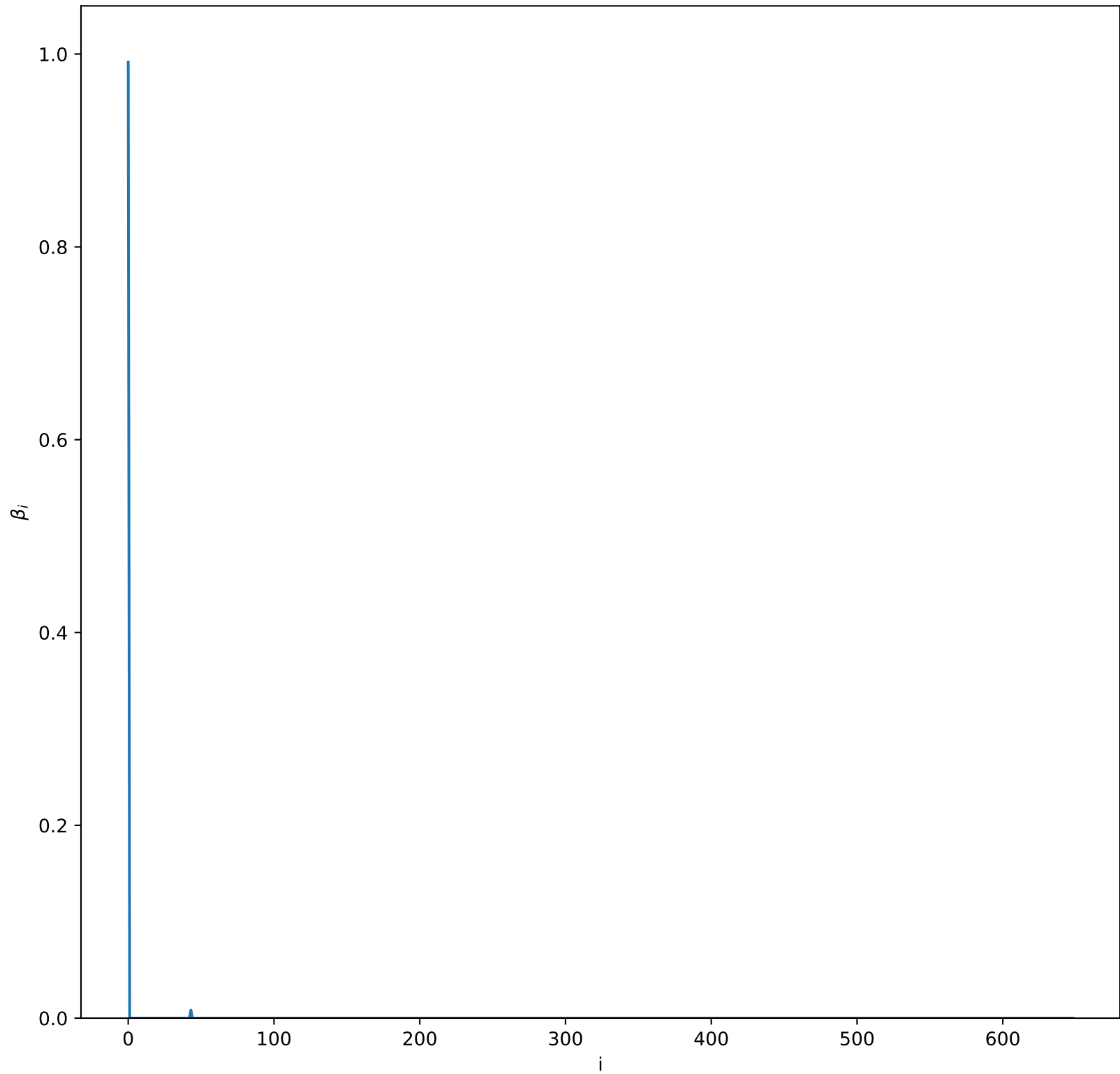
$\mu = 2.99$



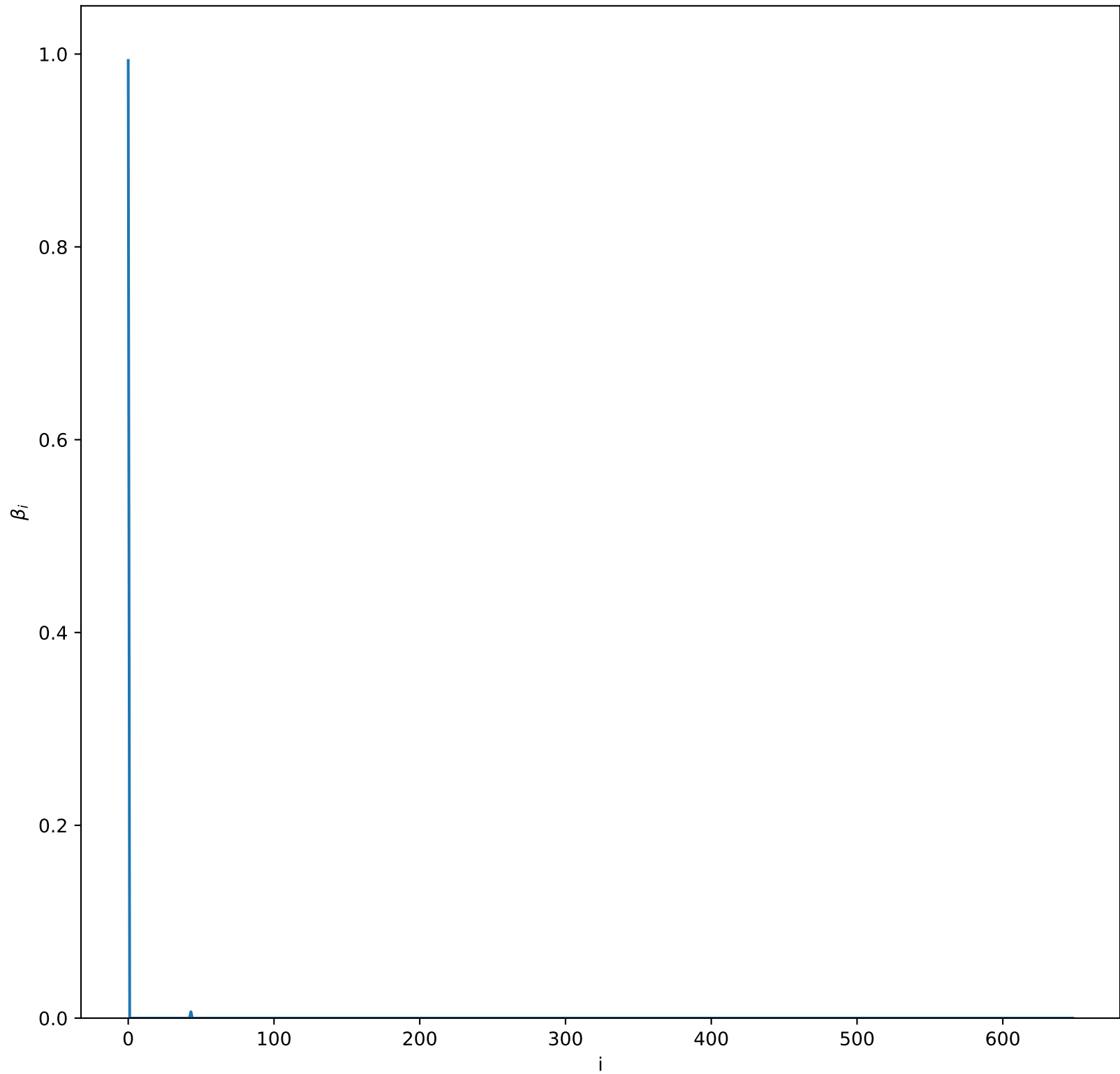
$\mu = 3.00$



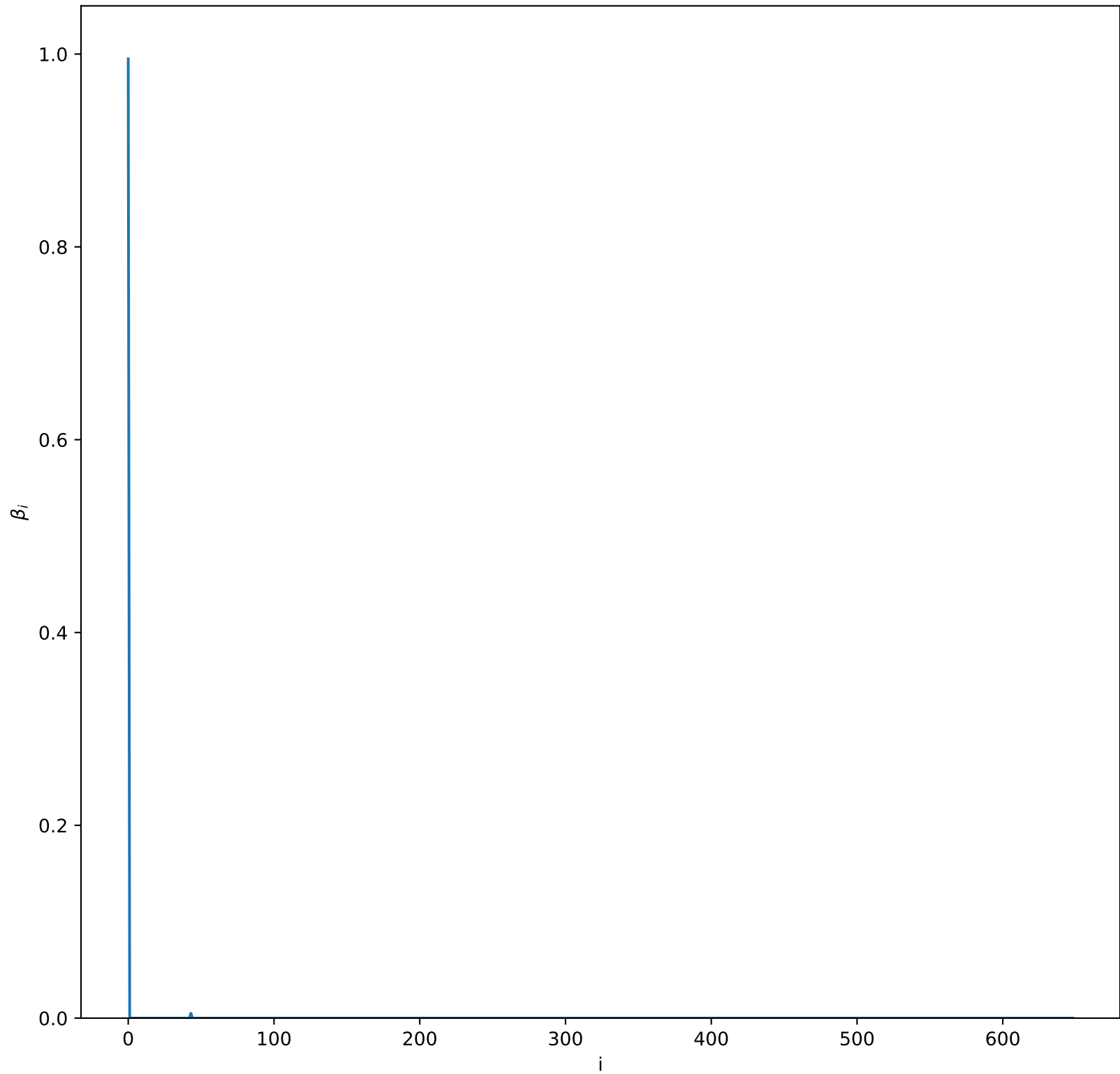
$\mu = 3.01$



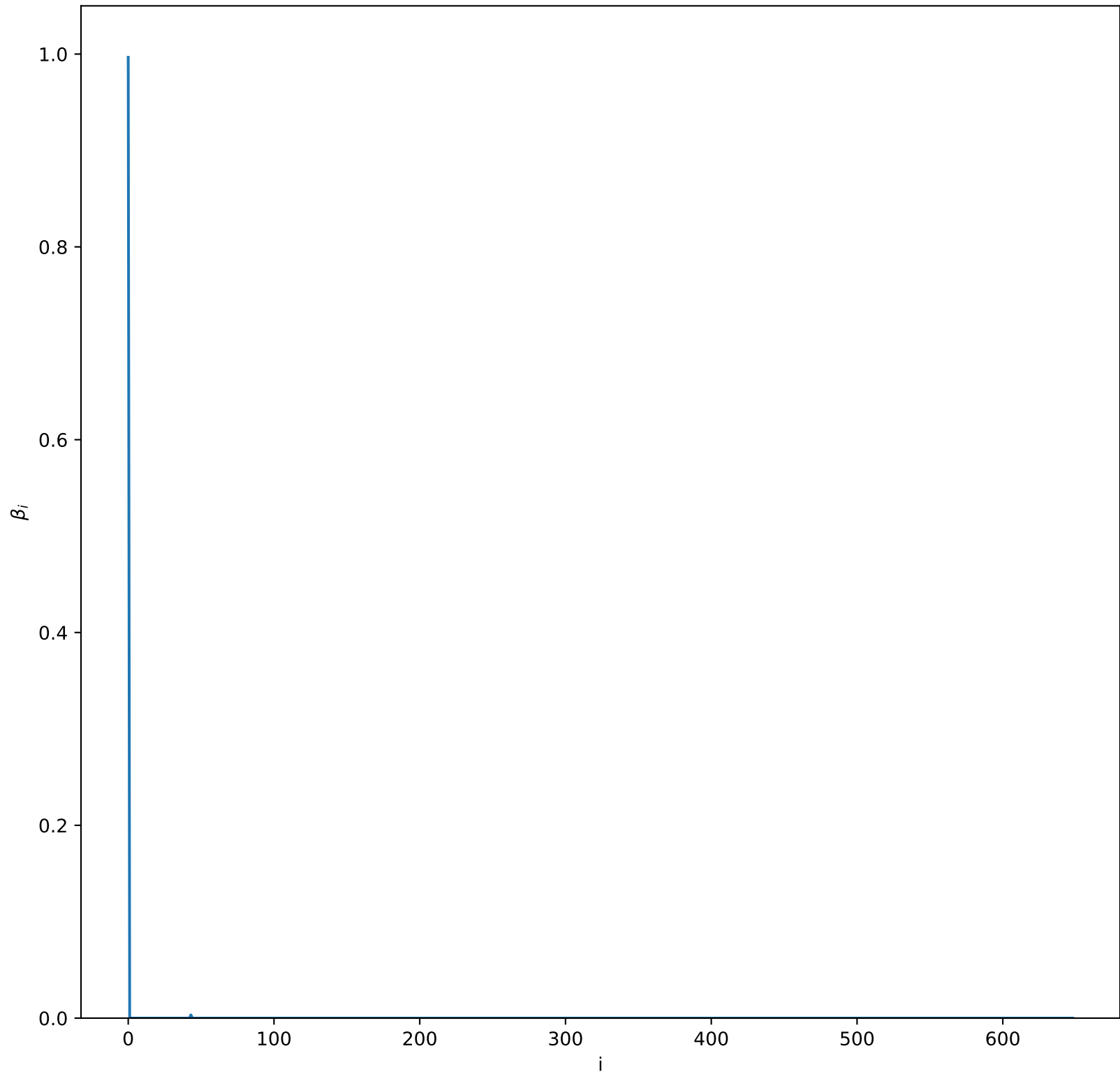
$\mu = 3.02$



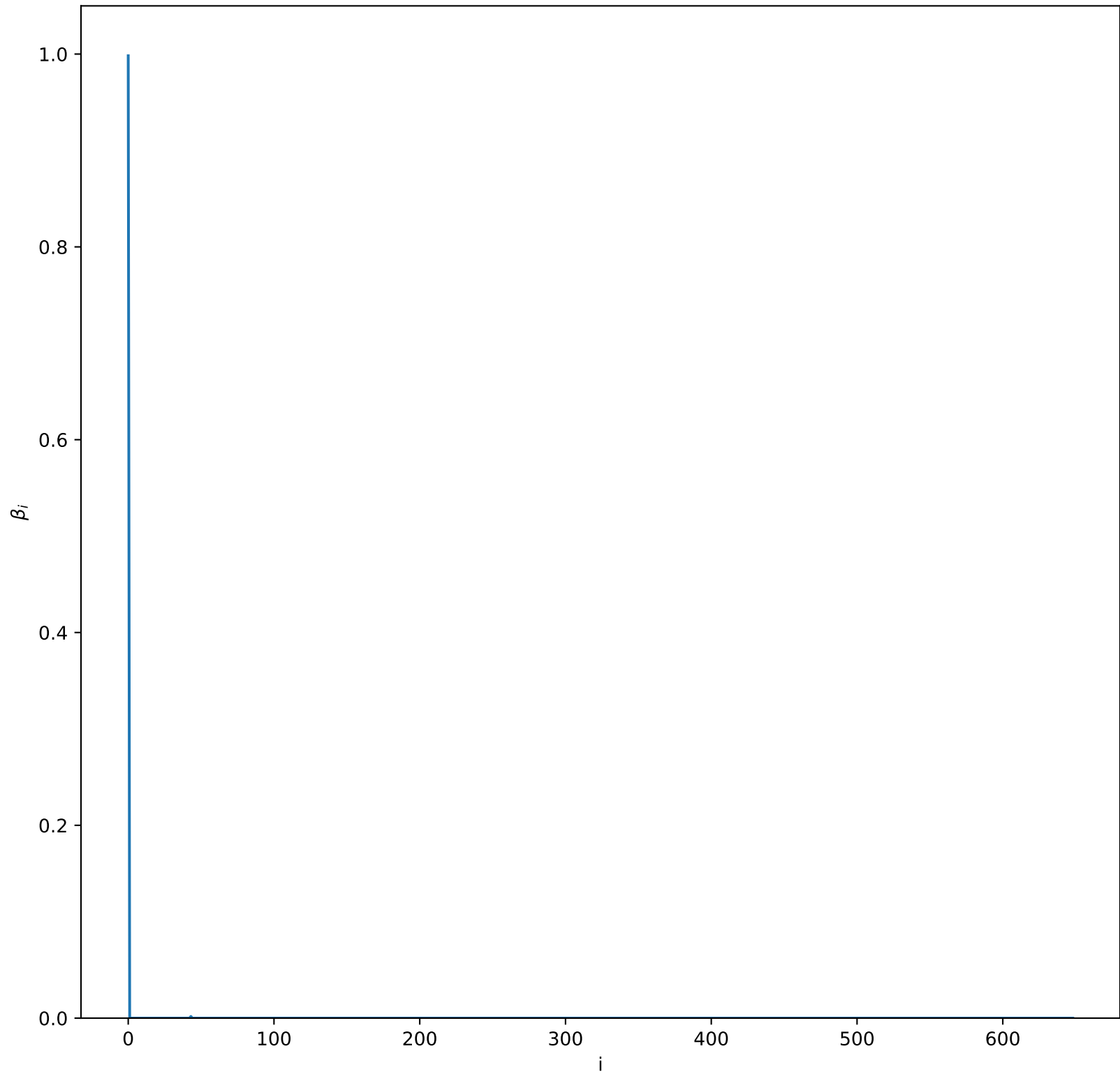
$\mu = 3.03$



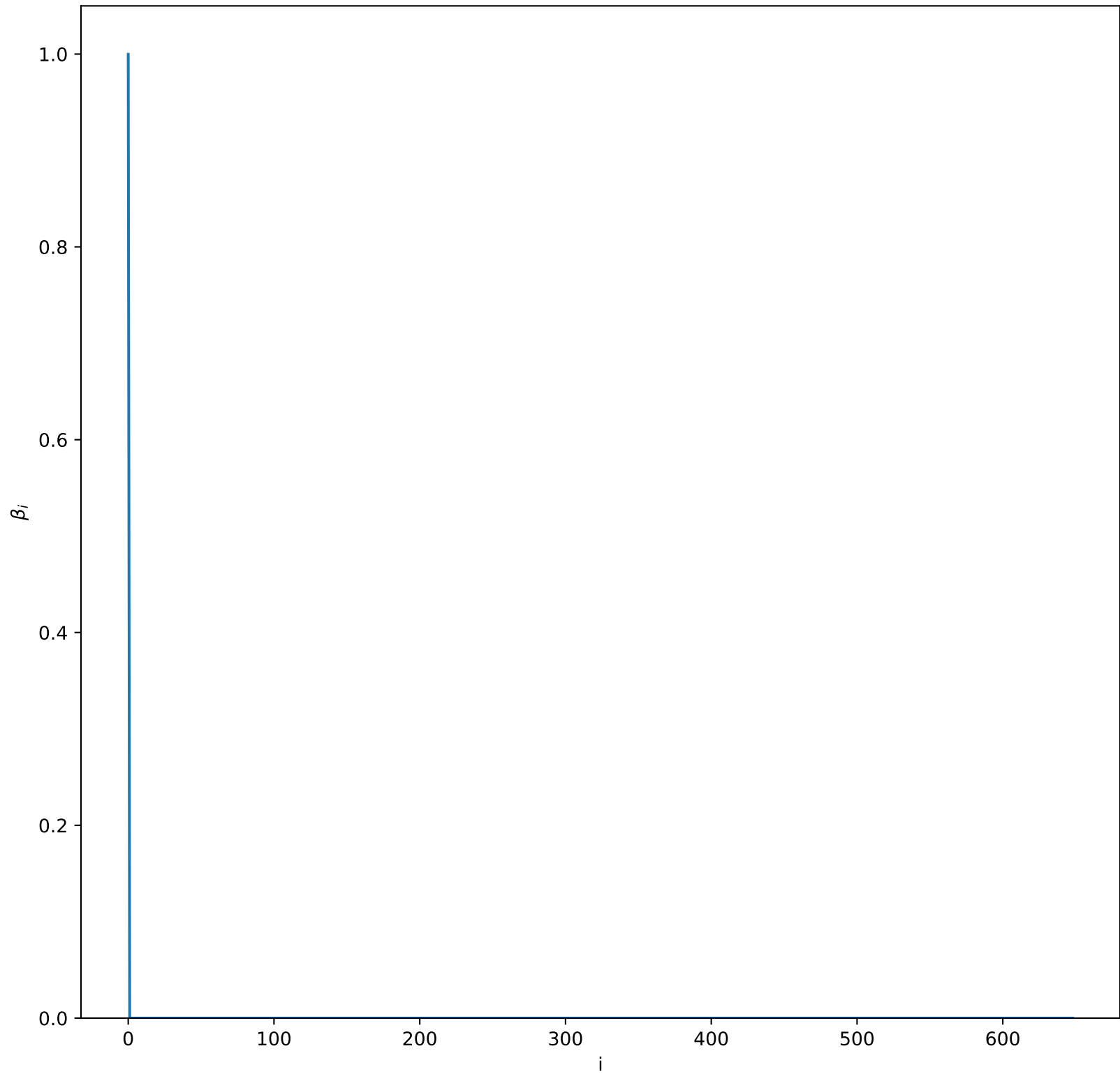
$\mu = 3.04$



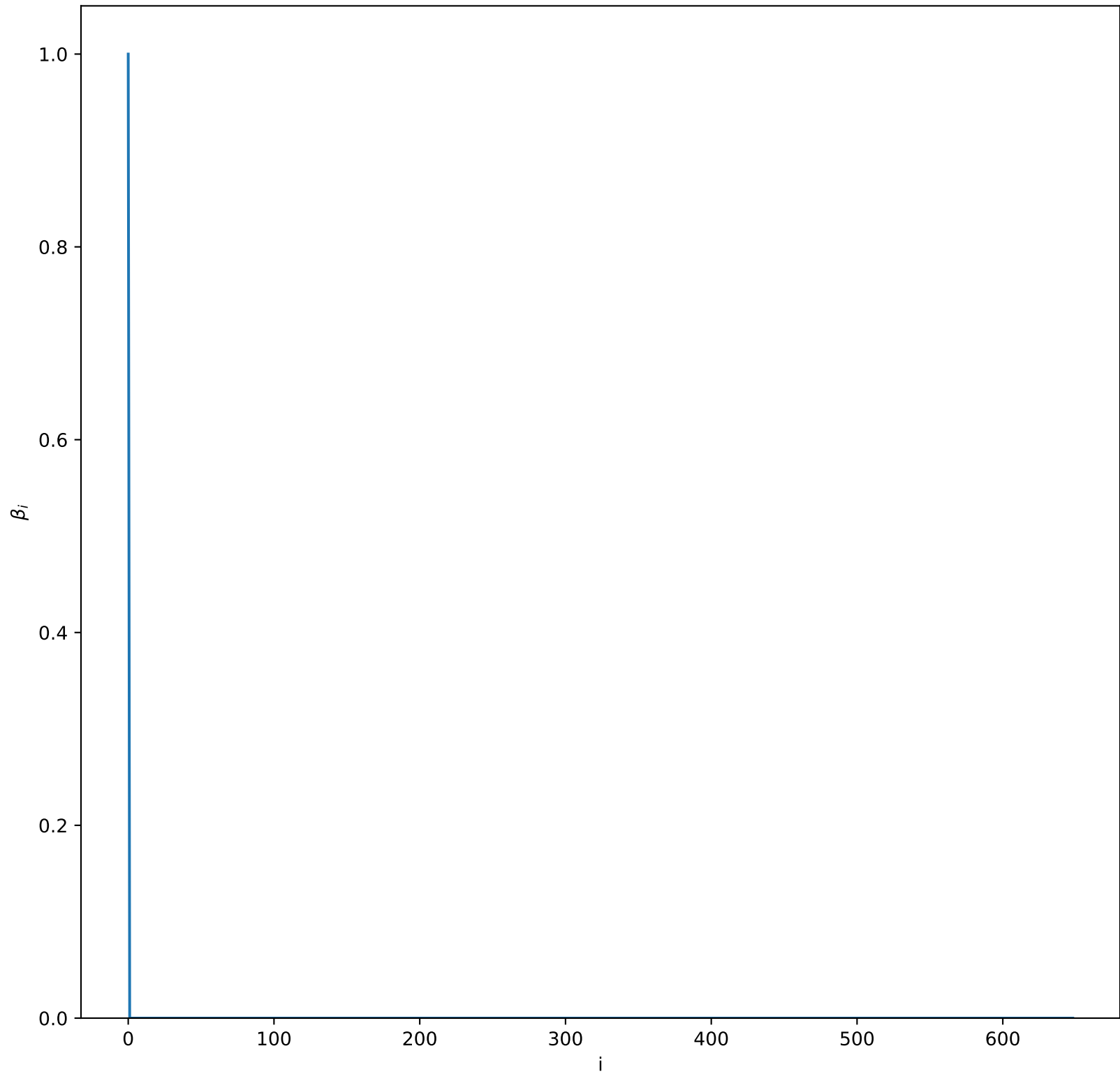
$\mu = 3.05$



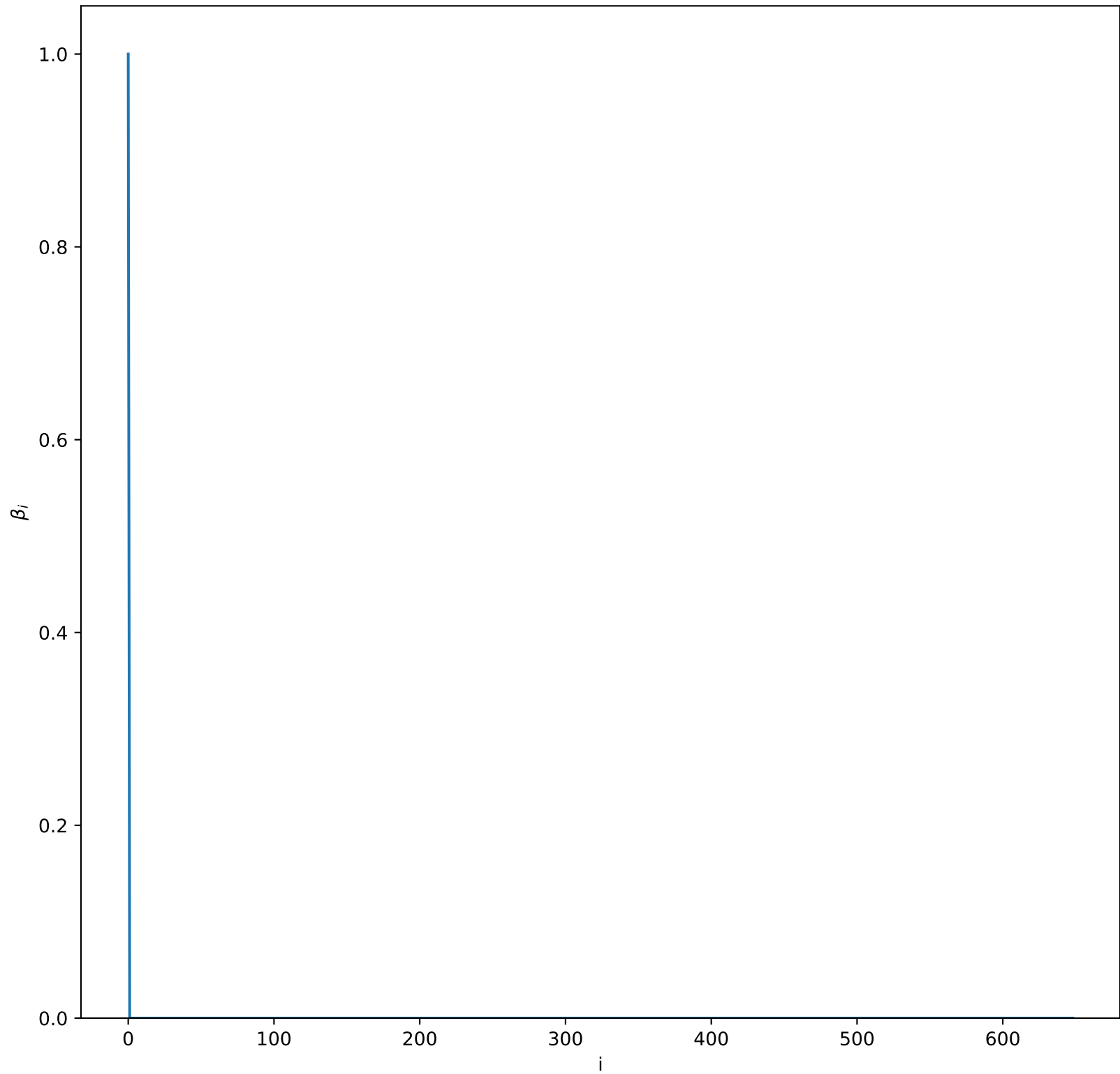
$\mu = 3.06$



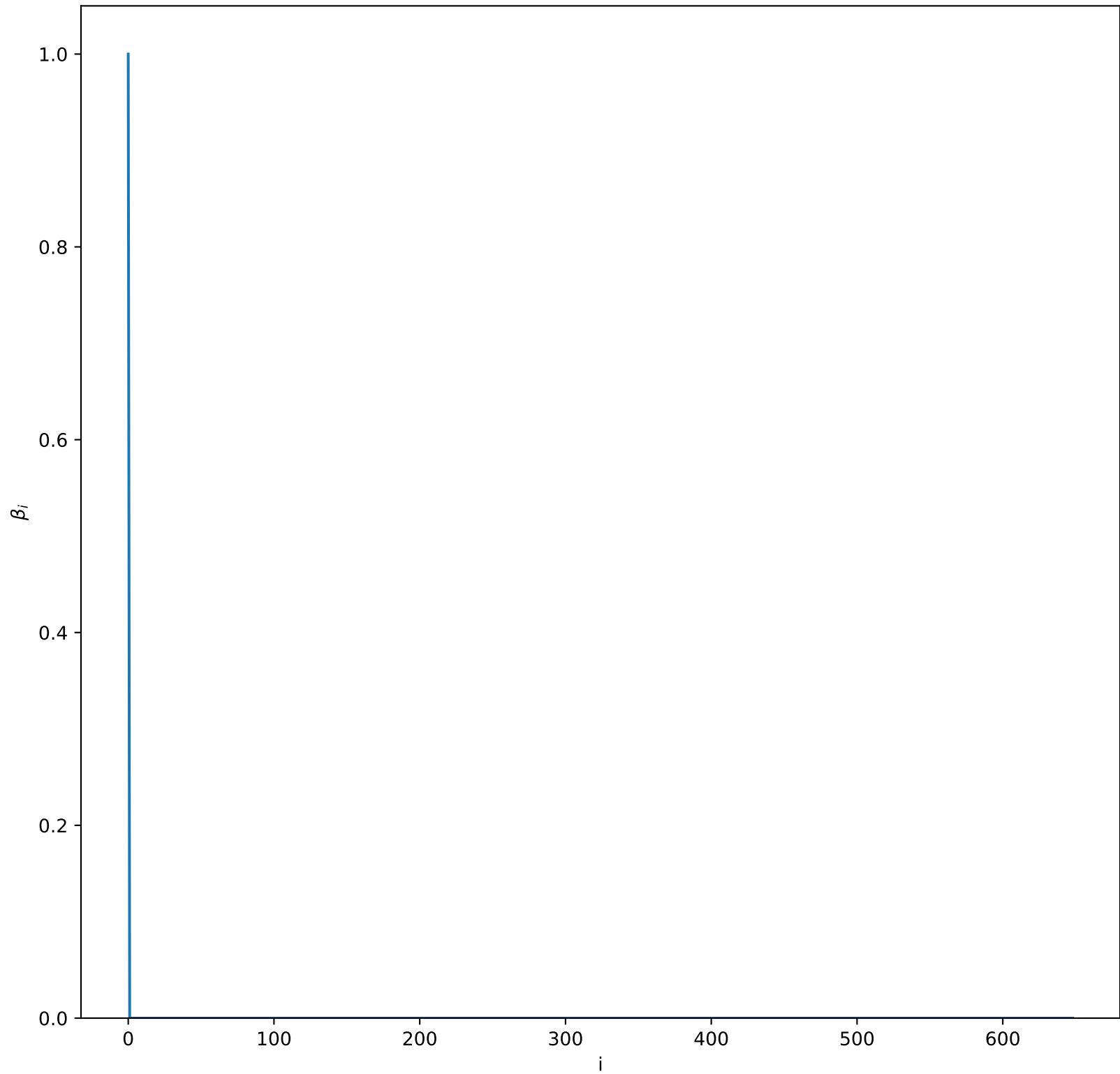
$\mu = 3.07$



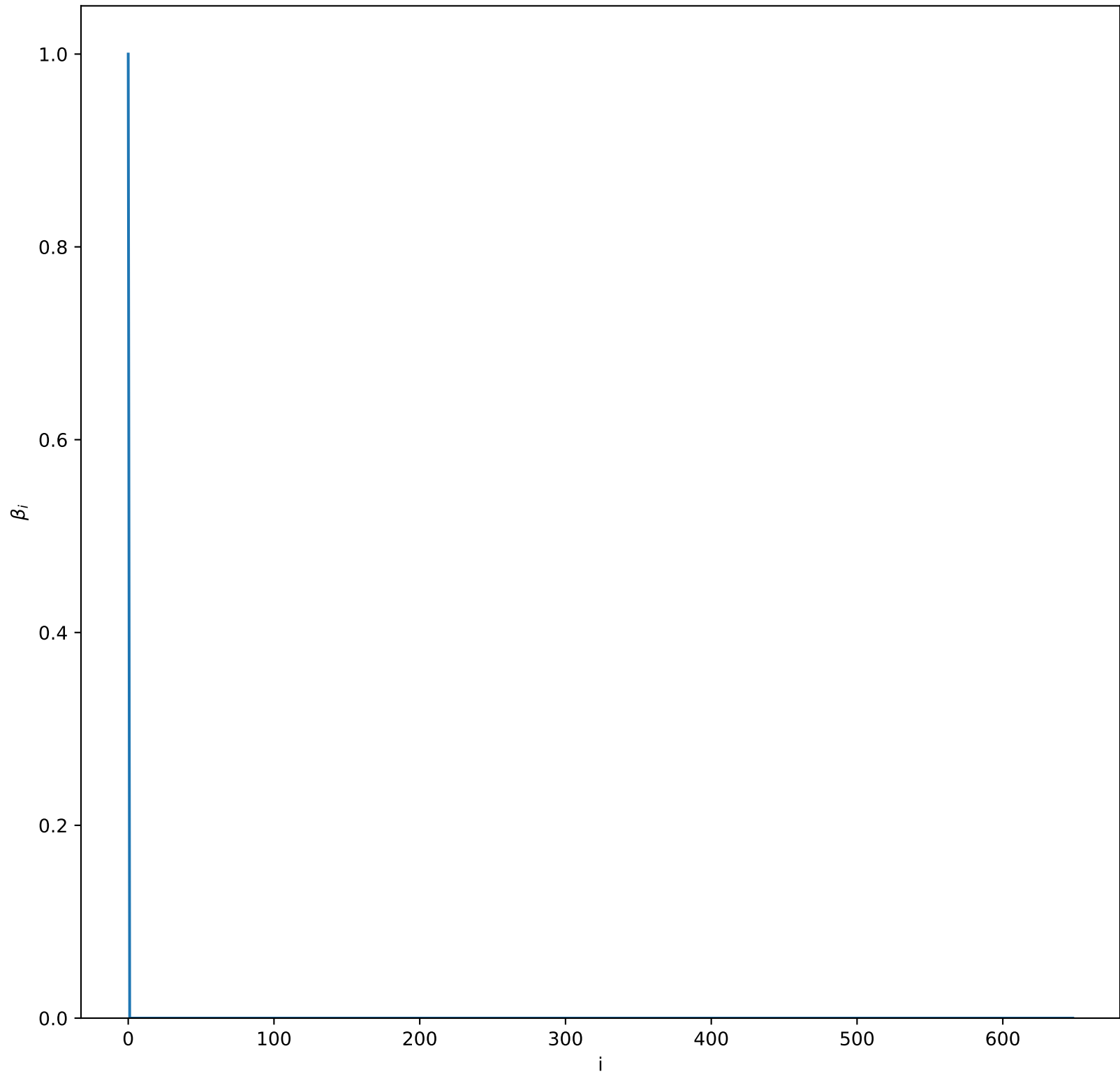
$\mu = 3.08$



$\mu = 3.09$



$\mu = 3.10$



Factor Search на основе технического анализа (Пугач, Морозов):

$$\begin{cases} \beta^T \underbrace{\Sigma}_{\text{}} \beta - \mu \underbrace{\bar{x}^T}_{\text{}} \beta + c (\mathbf{y} - \mathbf{X}^T \beta)^T (\mathbf{y} - \mathbf{X}^T \beta) \rightarrow \min(\beta) \\ \mathbf{1}^T \beta = 1; \quad \beta \geq \mathbf{0} \end{cases} \quad (6)$$

Наш Factor Search с учетом экспертного мнения:

$$\begin{cases} \beta^T \left(\underbrace{(1-\alpha)\Sigma + \alpha\mathbf{B}}_{\text{}} \right) \beta - \mu \left(\underbrace{(1-\alpha)\bar{x} + \alpha z}_{\text{}} \right)^T \beta \\ \quad + c (\mathbf{y} - \mathbf{X}^T \beta)^T (\mathbf{y} - \mathbf{X}^T \beta) \rightarrow \min(\beta) \\ \mathbf{1}^T \beta = 1; \quad \beta \geq \mathbf{0} \end{cases} \quad (7)$$

$$\begin{cases} \beta^T \left((1 - \alpha) \Sigma + \alpha \mathbf{B} \right) \beta - \mu \left((1 - \alpha) \bar{\mathbf{x}} + \alpha \mathbf{z} \right)^T \beta + \\ \quad + c (\mathbf{y} - \mathbf{X}^T \beta)^T (\mathbf{y} - \mathbf{X}^T \beta) \rightarrow \min(\beta) \\ \mathbf{1}^T \beta = 1; \quad \beta \geq \mathbf{0} \end{cases} \quad (8)$$

Если $\mathbf{y} = (y_t, t = 1, \dots, T)$ — анализируемый портфель, то всякая комбинация параметров (α, μ, c) дает решение этой задачи как совокупность коэффициентов $\hat{\beta} = (\beta_i, i = 1, \dots, n)$ и следовательно состав портфеля:

$\hat{\mathbb{I}} = \{i : \beta_i > 0\} \in \mathbb{I}$, $\hat{n} = |\hat{\mathbb{I}}|$ — размер портфеля

Это ничто иное как параметры регуляризации модели.

Для выбора их значений для данного наблюдаемого сигнала \mathbf{y} необходимо использовать некоторый критерий верификации модели (обобщающей способности). Например Leave One Out (LOO).

Критерий Leave One Out для Factor Search

Итак, пусть $\mathbf{y} = (y_t, t = 1, \dots, T)$ — анализируемый портфель. В предположении, что администратор построил его по принципу рассмотренному ранее, остается найти значения параметров α, μ, c

Критерий LOO дает количественную оценку всякому варианту α, μ, c .

$$LOO(\alpha, \mu, c) = \frac{1}{T} \sum_{t=1}^T \left(y_t - \sum_{i=1}^n \beta_{\alpha, \mu, c, i}^{(t)} x_{t,i} \right)^2 \rightarrow \min(\alpha, \mu, c) \quad (9)$$

Одна трудность — вычисление $\beta^{(t)}$ с одним пропущенным наблюдением нужно повторить очень много раз (T раз).

Для преодоления этой проблемы Пугач и Морозов разработали специальную форму критерия LOO, учитывающую специфику задачи Factor Search.

Критерий Leave One Out для Factor Search

Специальная форма критерия LOO для Factor Search:

$$LOO(\alpha, \mu, c) = \frac{1}{T} \sum_{t=1}^T \left(\frac{y_t - \sum_{i=1}^n \beta_{\alpha, \mu, c, i} x_{t, i}}{1 - \mathbf{x}_{\mathbb{I}_{\alpha, \mu, c}, t}^T \left(\mathbf{X}_{\mathbb{I}_{\alpha, \mu, c}} \mathbf{X}_{\mathbb{I}_{\alpha, \mu, c}}^T \right)^{-1} \mathbf{x}_{\mathbb{I}_{\alpha, \mu, c}, t}} \right)^2 \rightarrow \min(\alpha,$$

Здесь для каждой комбинации набора (α, μ, c) достаточно один раз найти вектор коэффициентов регрессии, обратив соответствующую матрицу.

$n = 650$; $T = 240$

Строим портфели:

$$\begin{cases} \beta_{\mu,\alpha}^* = \arg \min (\alpha \beta^T \Sigma \beta + (1 - \alpha) \beta^T \mathbf{B} \beta + \mu \bar{\mathbf{x}}^T \beta) \\ \mathbf{1}^T \beta = 1 \\ \beta \geq \mathbf{0} \end{cases} \quad (10)$$

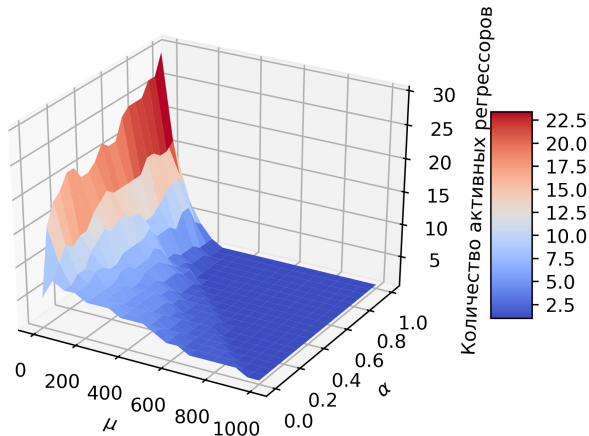
Множество активных индексов: $\mathbb{I}_{\mu,\alpha}^* = \{i : \beta_{\mu,\alpha,i}^* > 0\}$

$$y_{\mu,\alpha,t} = \sum_{i \in \mathbb{I}_{\mu,\alpha}^*} \beta_{\mu,\alpha,i}^* x_{t,i} + \xi_t \quad (11)$$

$\sigma^2(\xi) = 10\%$

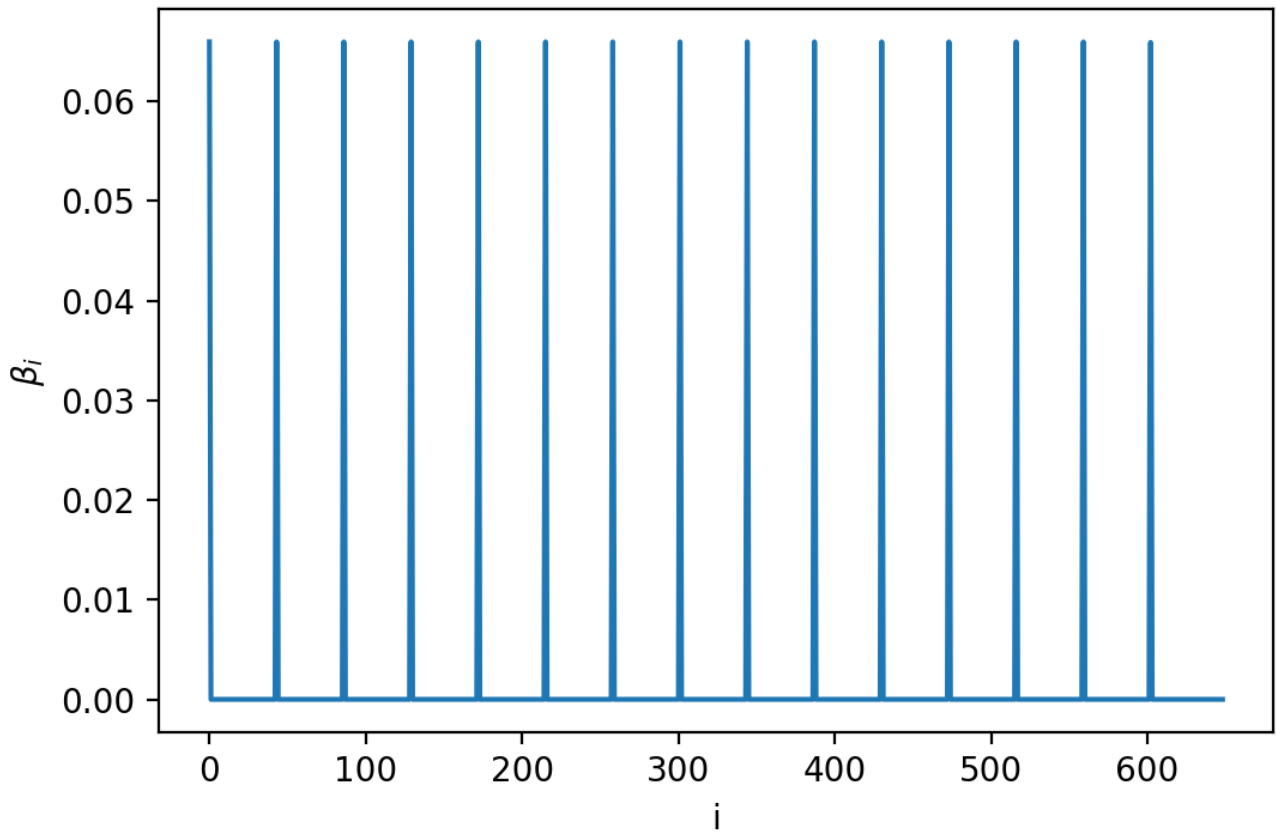
Экспериментальная часть

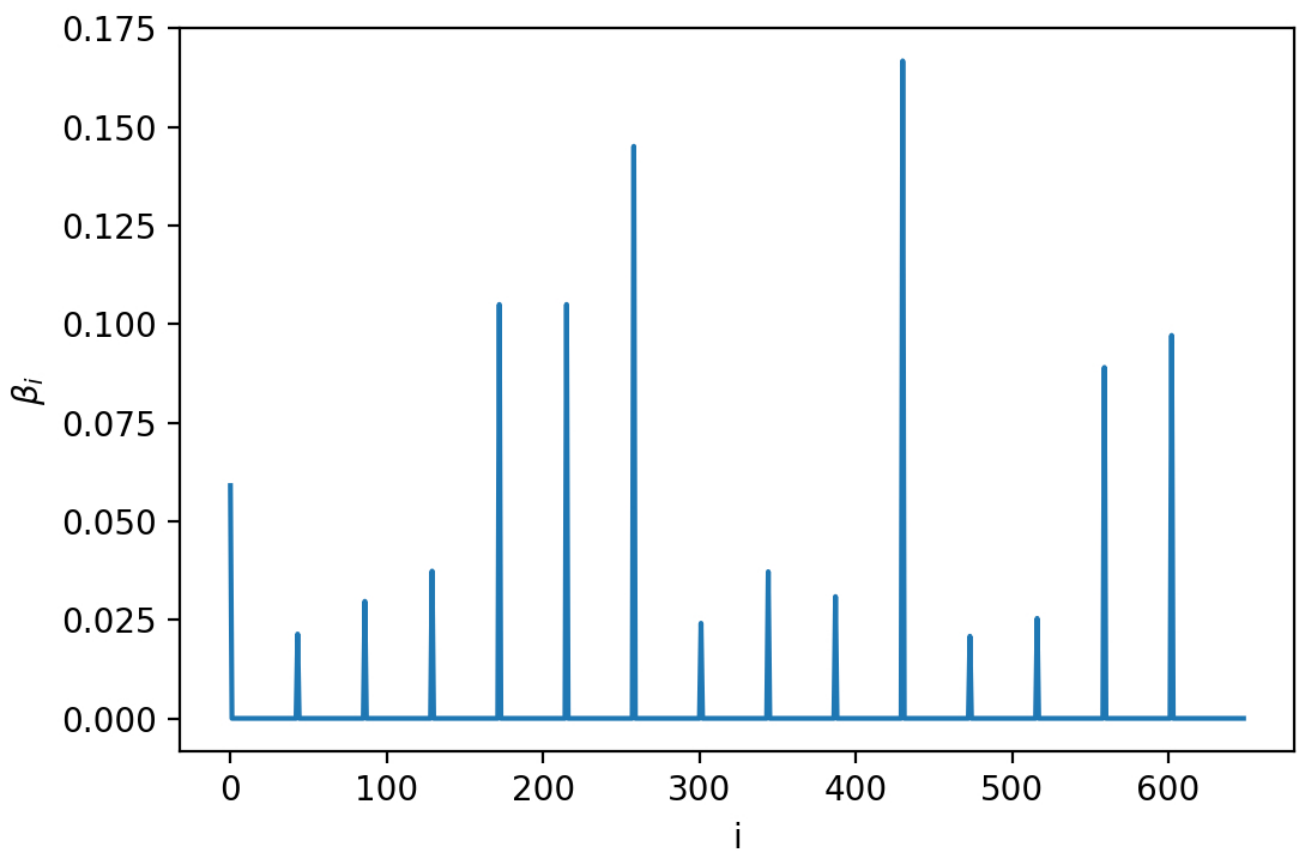
Зависимость размера портфеля : $|\mathbb{I}_{\mu, \alpha}^*|$ от α , μ :

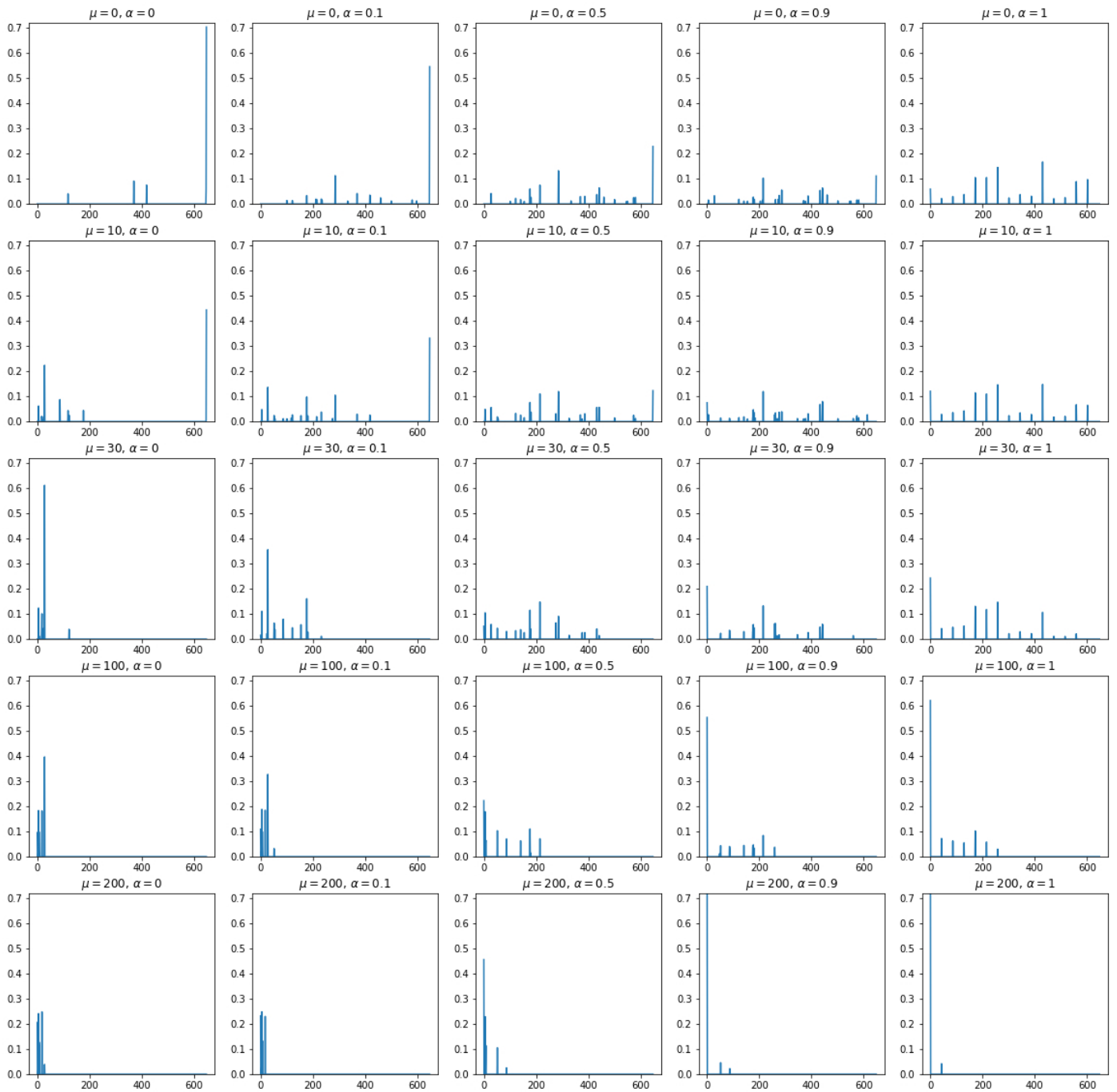


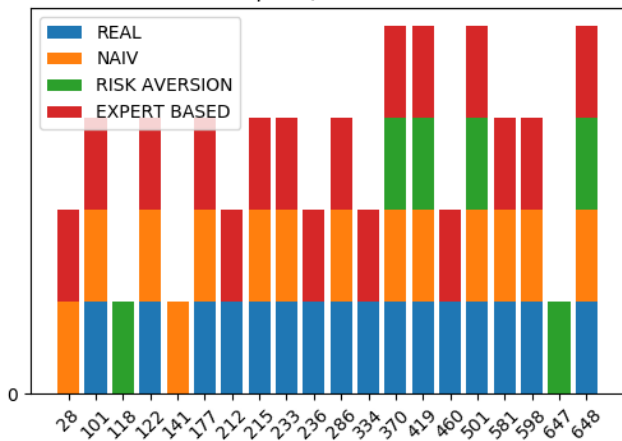
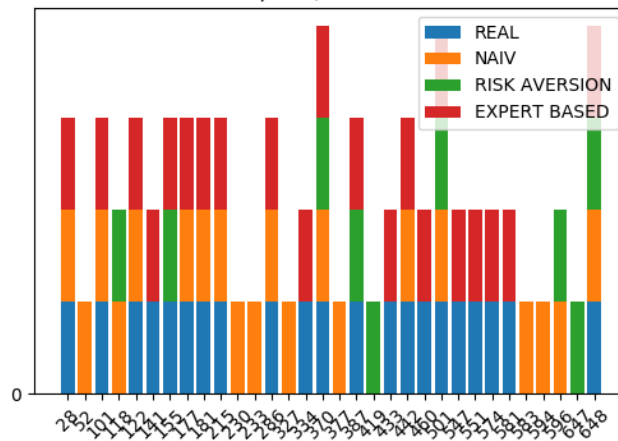
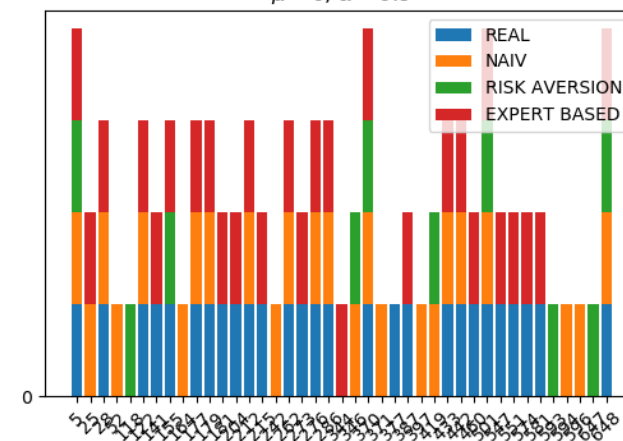
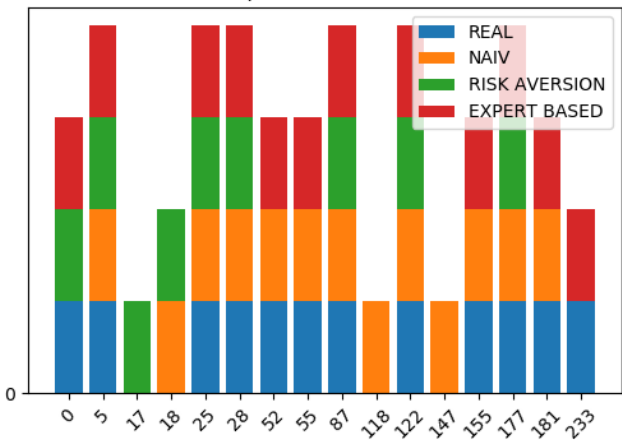
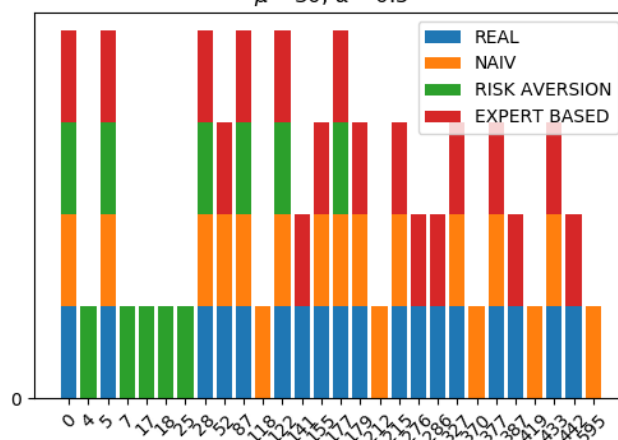
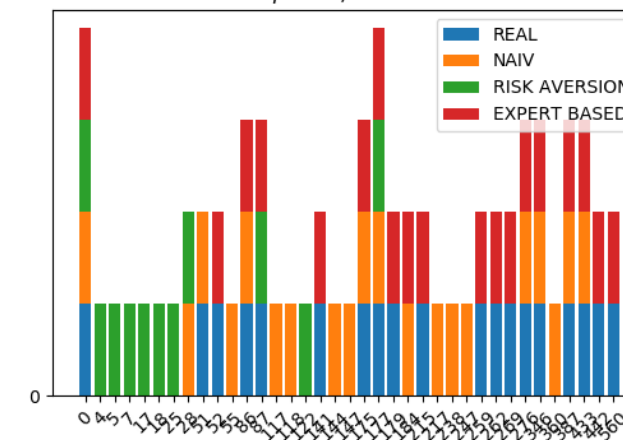
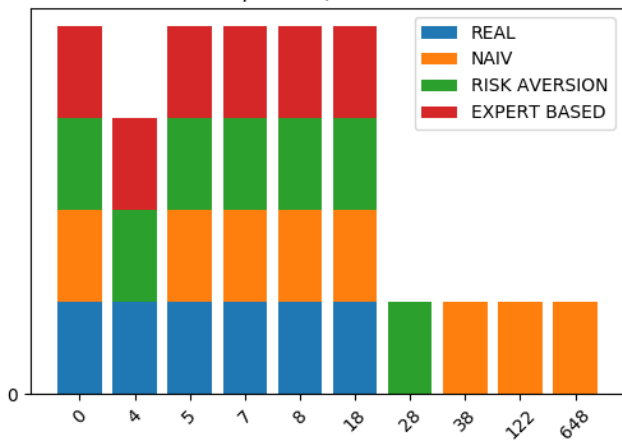
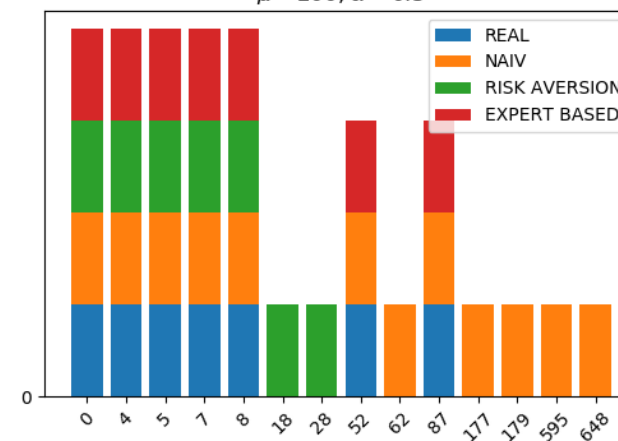
Спасибо за внимание!

$$\left\{ \begin{array}{l} \beta^T \left((1 - \alpha) \Sigma + \alpha \mathbf{B} \right) \beta - \mu \left((1 - \alpha) \bar{\mathbf{x}} + \alpha \mathbf{z} \right)^T \beta + \\ \quad + c (\mathbf{y} - \mathbf{X}^T \beta)^T (\mathbf{y} - \mathbf{X}^T \beta) \rightarrow \min(\beta) \\ \mathbf{1}^T \beta = 1; \\ \beta \geq \mathbf{0} \end{array} \right. \quad (12)$$







$\mu = 0, \alpha = 0.1$  $\mu = 0, \alpha = 0.5$  $\mu = 0, \alpha = 0.9$  $\mu = 30, \alpha = 0.1$  $\mu = 30, \alpha = 0.5$  $\mu = 30, \alpha = 0.9$  $\mu = 200, \alpha = 0.1$  $\mu = 200, \alpha = 0.5$  $\mu = 200, \alpha = 0.9$ 