

*Inconsistency of the MLE for the joint distribution
of interval censored survival times
and continuous marks*

Marloes Maathuis
University of Washington

joint work with
Jon Wellner, Michael Hudgens and Peter Gilbert

Motivation: VAX004 trial

- Phase III randomized, placebo-controlled HIV/AIDS vaccine trial

Motivation: VAX004 trial

- Phase III randomized, placebo-controlled HIV/AIDS vaccine trial
- Goals:
 - Evaluate vaccine efficacy
 - Evaluate dependence of vaccine efficacy on genetic sequence of infecting virus - does the vaccine protect better against viruses that are similar to the virus in the vaccine?

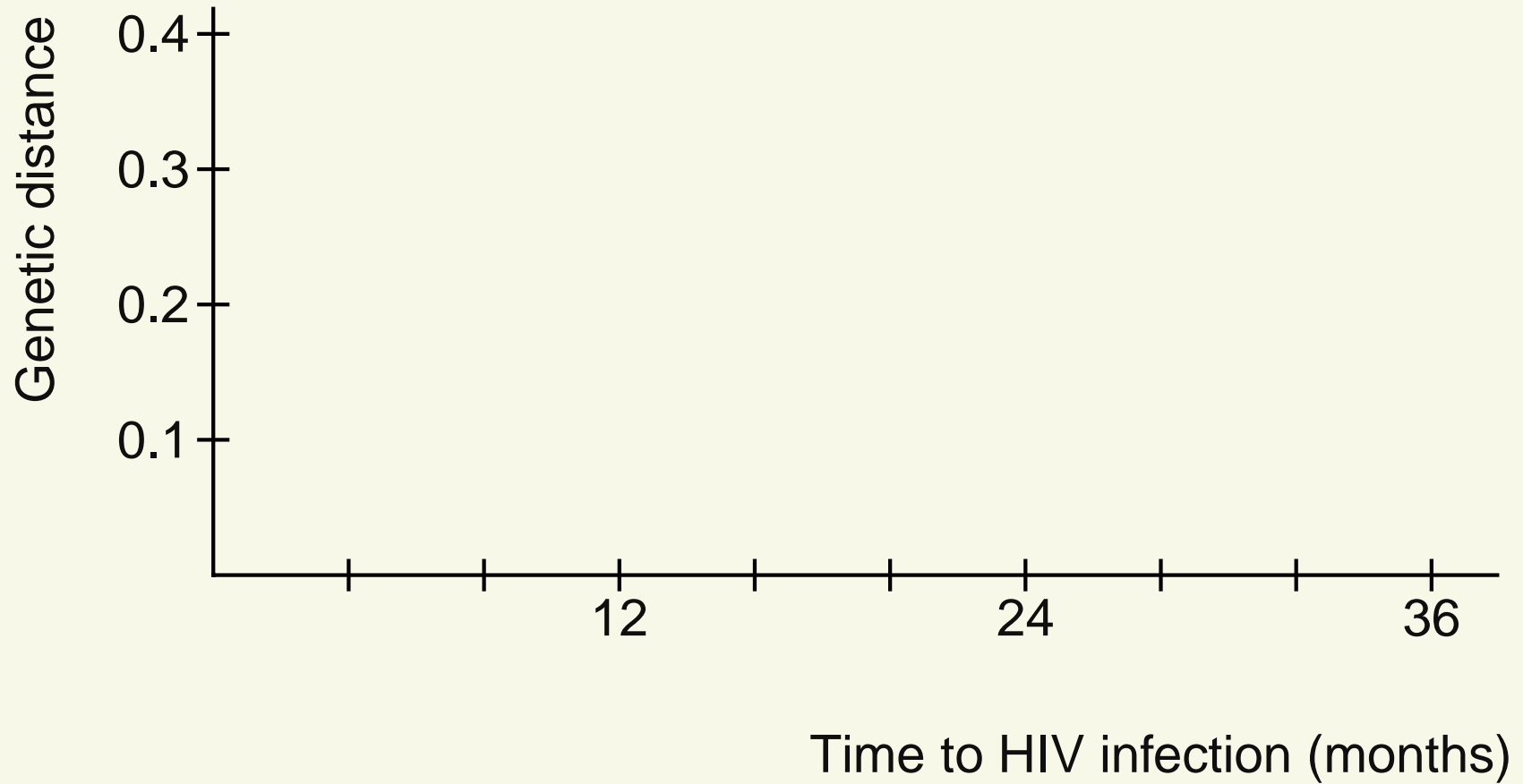
Motivation: VAX004 trial

- Phase III randomized, placebo-controlled HIV/AIDS vaccine trial
- Goals:
 - Evaluate vaccine efficacy
 - Evaluate dependence of vaccine efficacy on genetic sequence of infecting virus - does the vaccine protect better against viruses that are similar to the virus in the vaccine?
- Interested in the joint distribution of:
 - X : time to HIV infection
 - Y : genetic distance between the infecting virus and the virus present in the vaccine. Natural to treat Y as continuous variable (Gilbert et al, 2001).

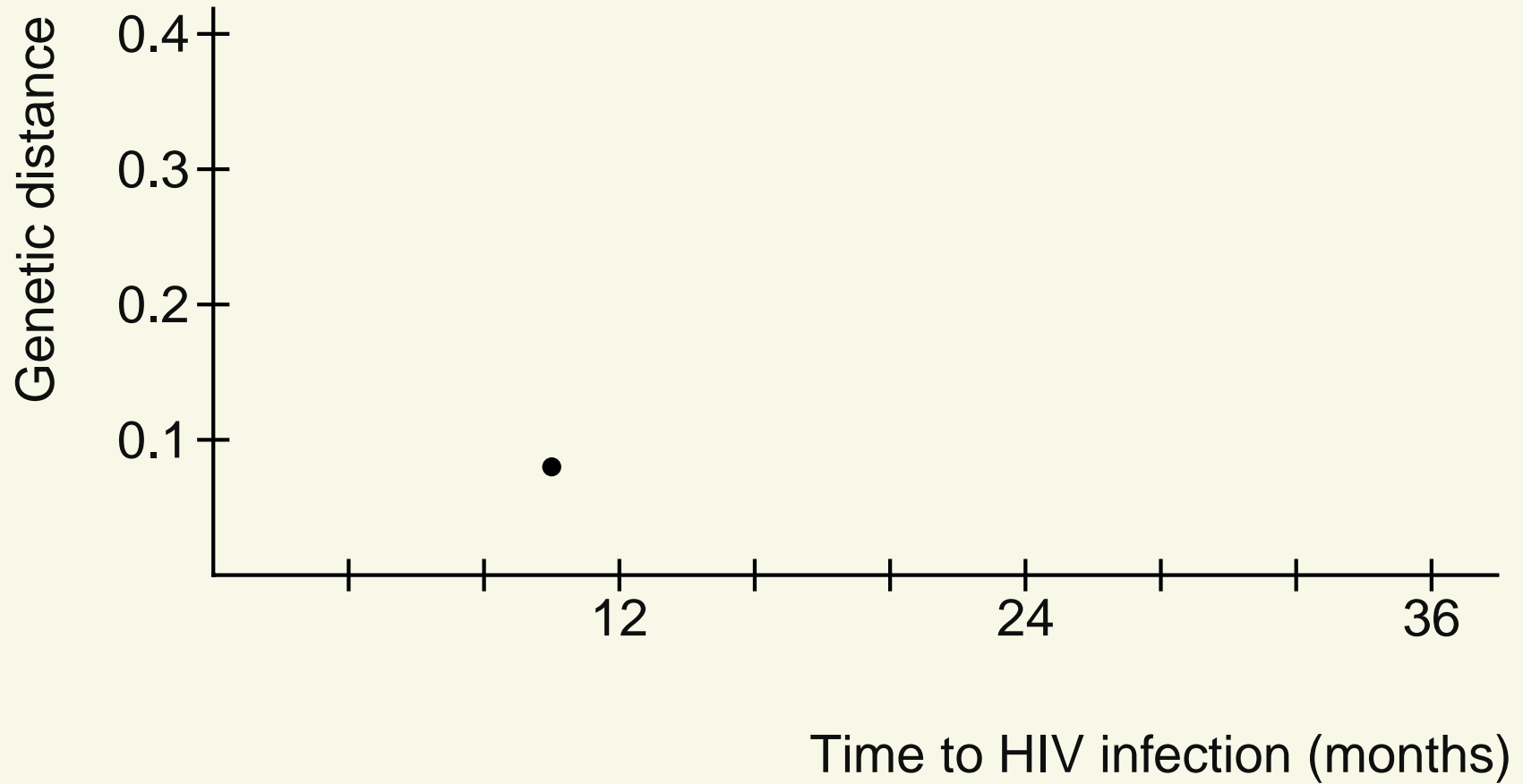
Motivation: VAX004 trial

- Phase III randomized, placebo-controlled HIV/AIDS vaccine trial
- Goals:
 - Evaluate vaccine efficacy
 - Evaluate dependence of vaccine efficacy on genetic sequence of infecting virus - does the vaccine protect better against viruses that are similar to the virus in the vaccine?
- Interested in the joint distribution of:
 - X : time to HIV infection
 - Y : genetic distance between the infecting virus and the virus present in the vaccine. Natural to treat Y as continuous variable (Gilbert et al, 2001).
- X and Y cannot be observed directly:
 - X is subject to interval censoring
 - Y is only observed if a person is infected: mark variable

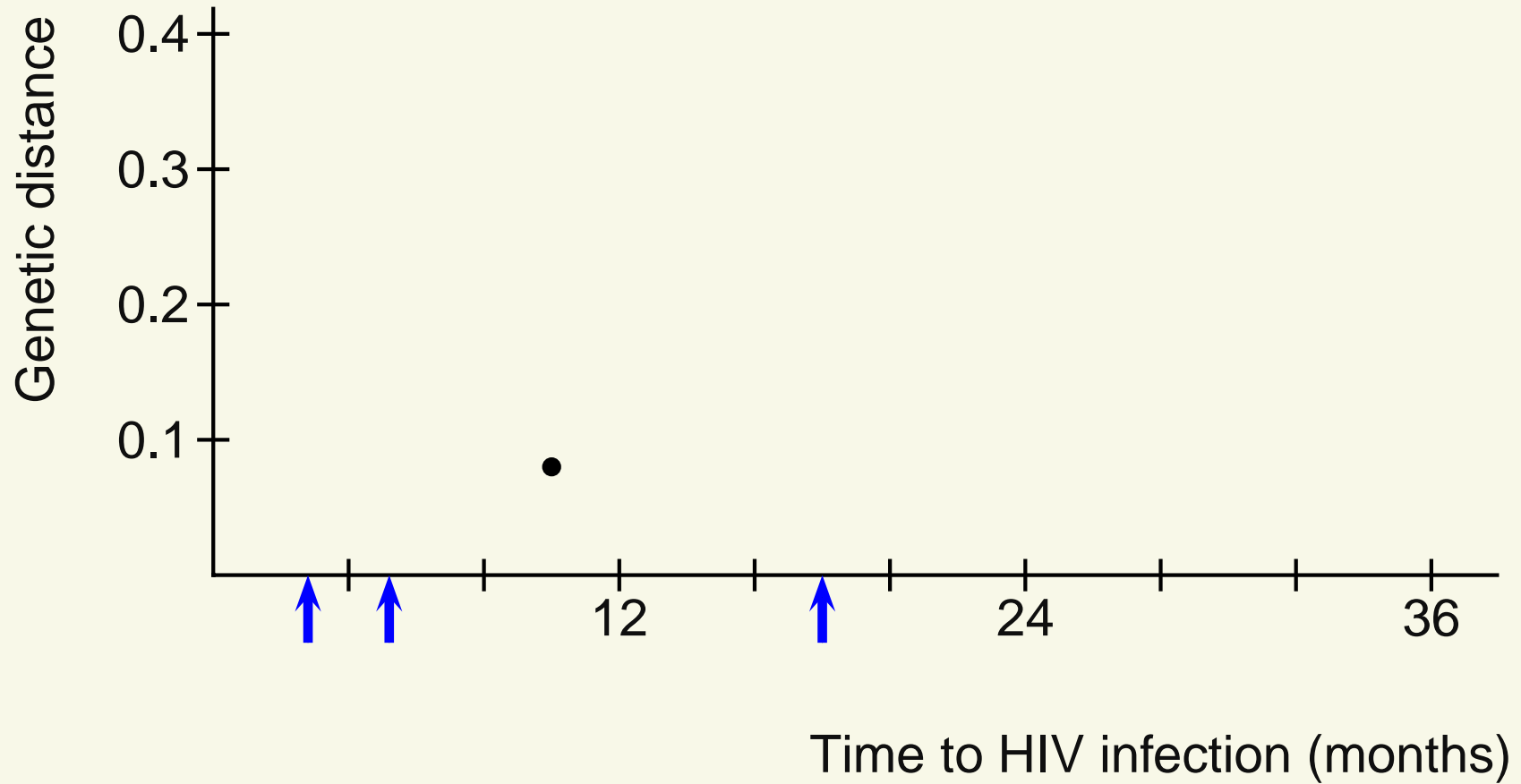
Observed sets



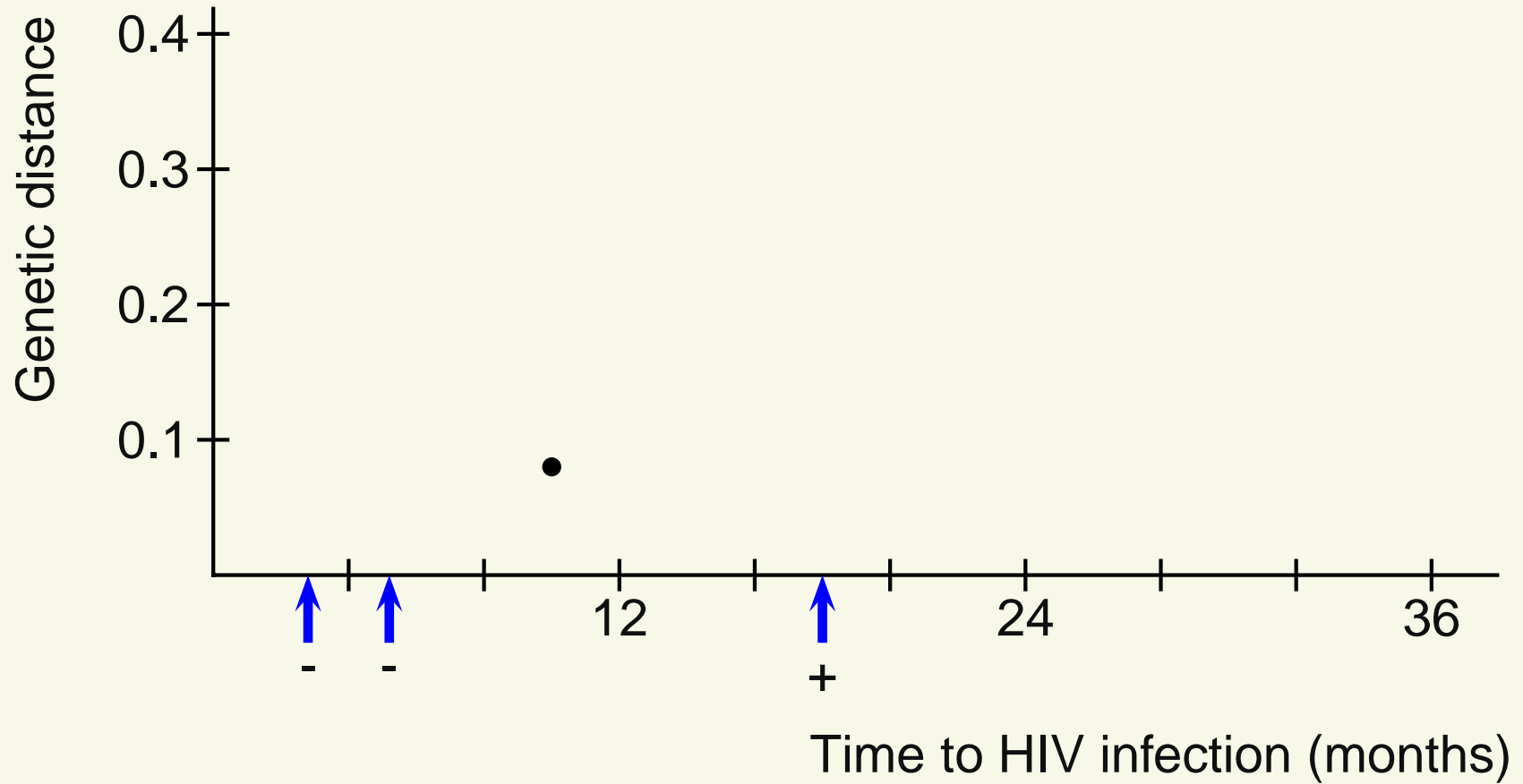
Observed sets



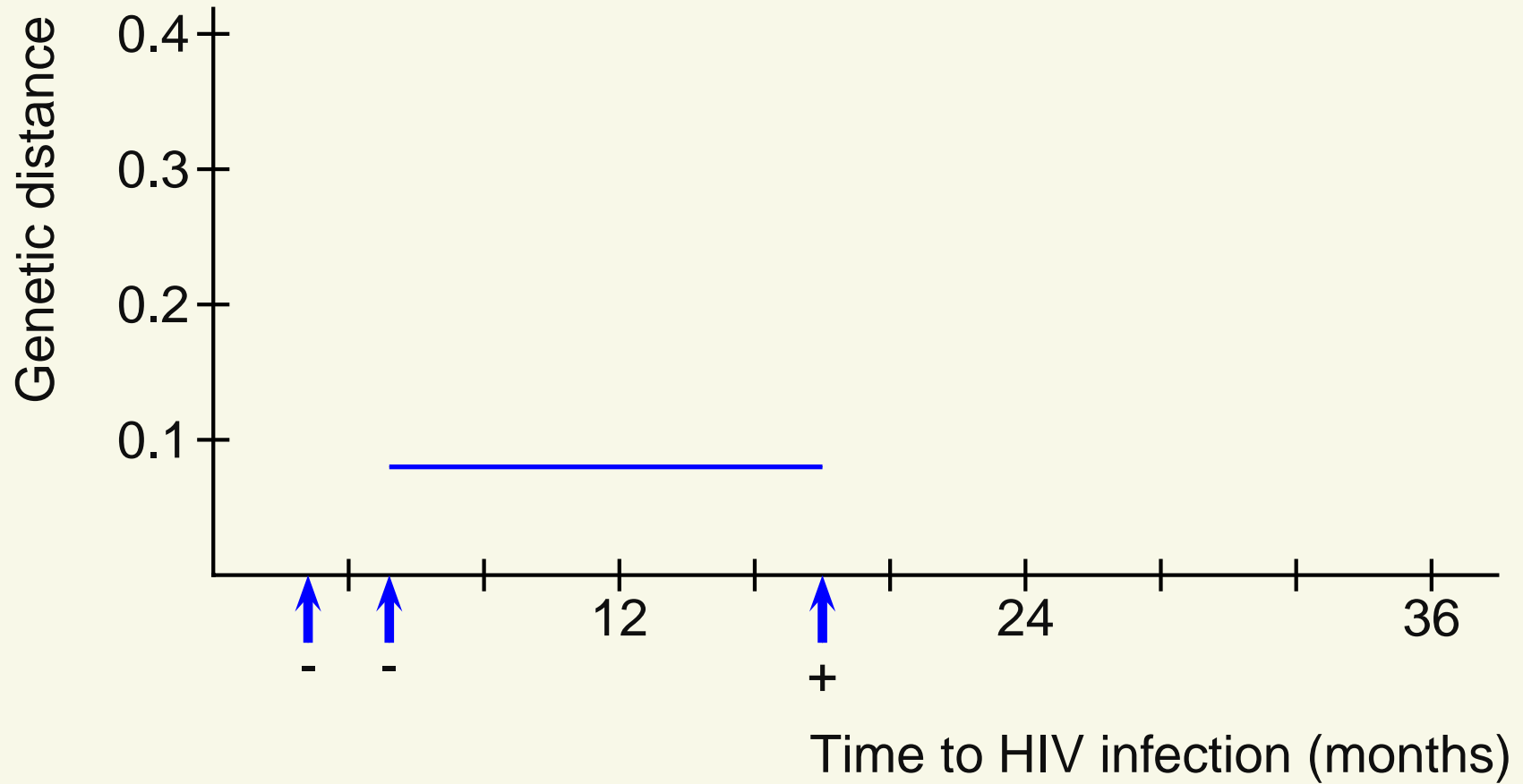
Observed sets



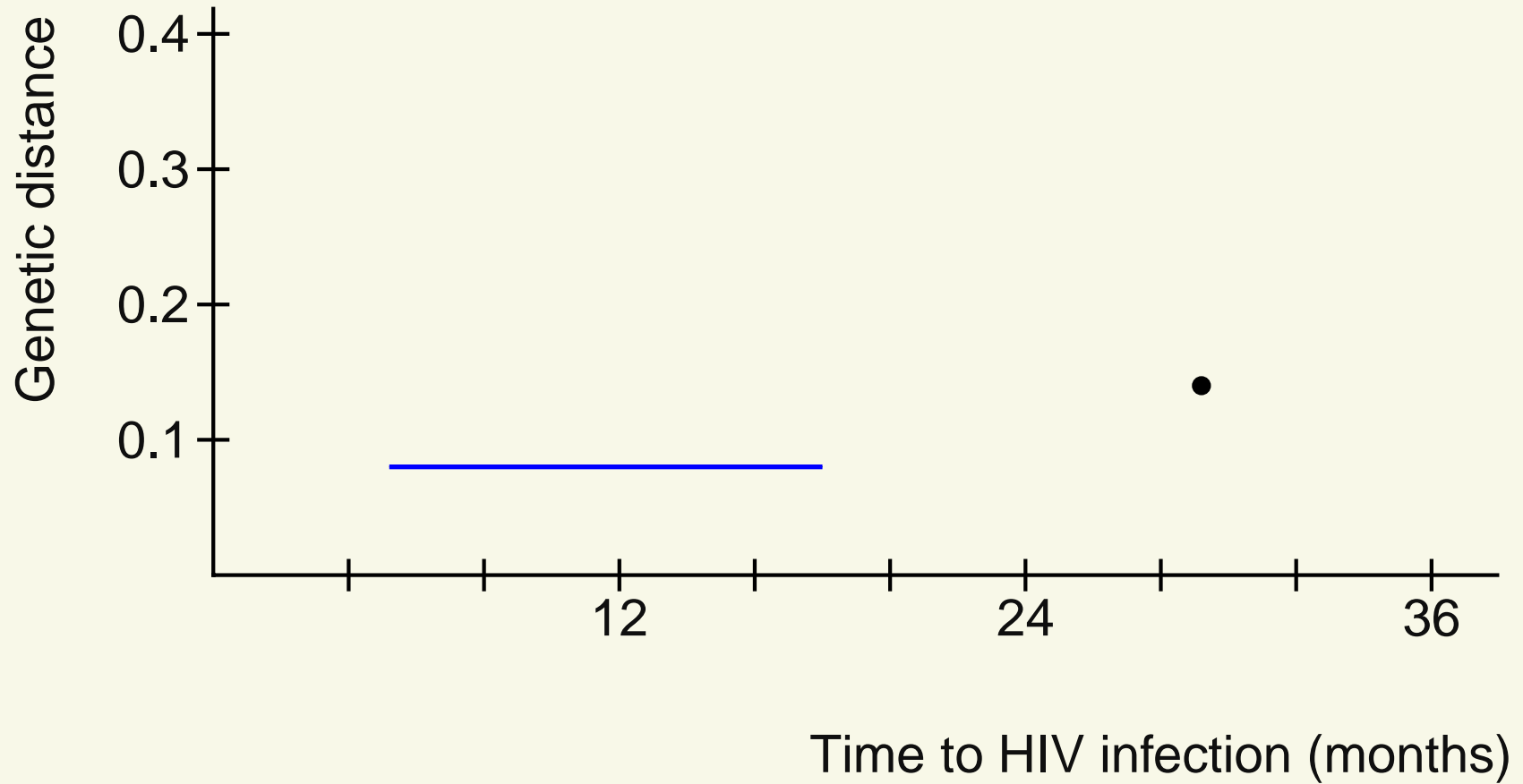
Observed sets



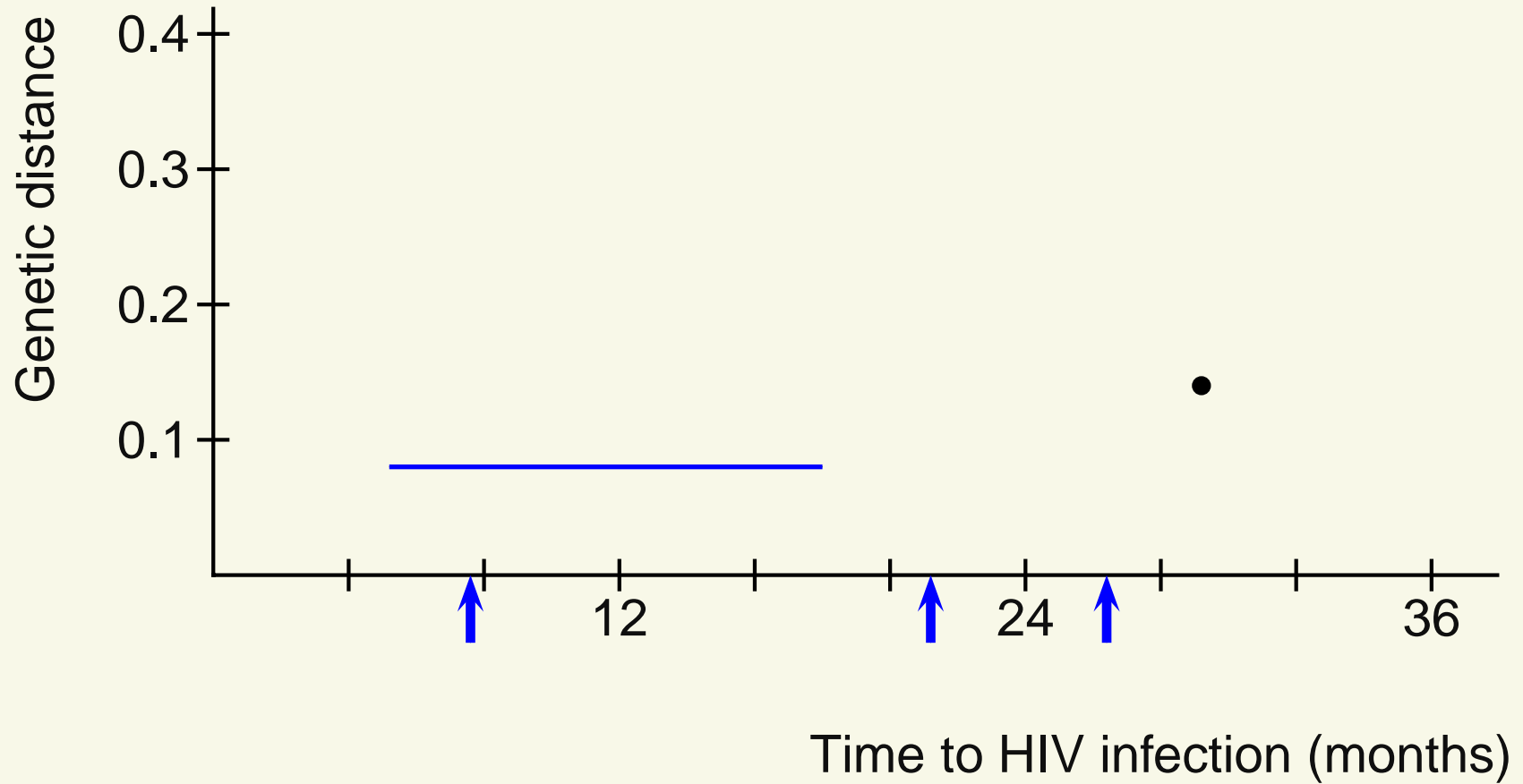
Observed sets



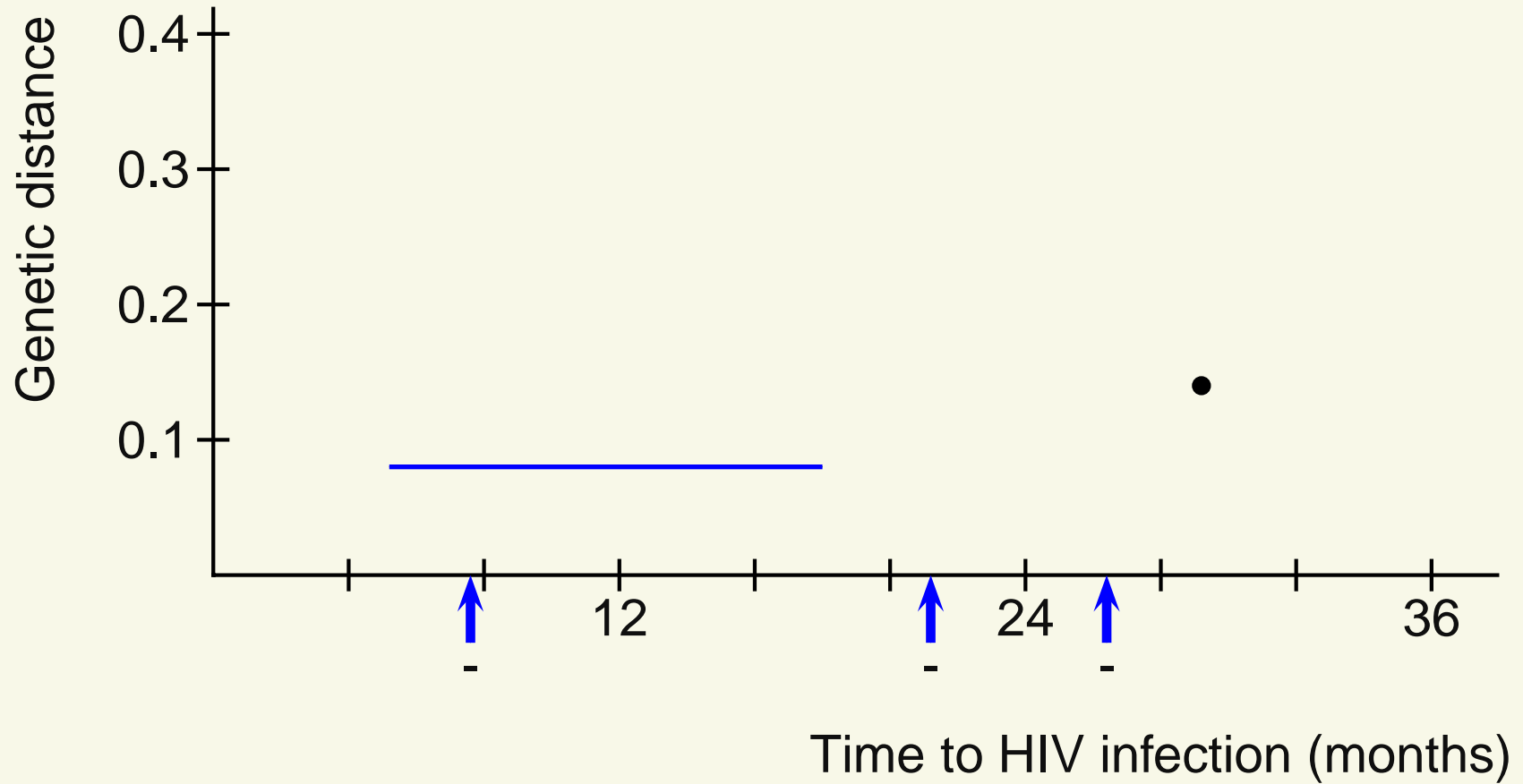
Observed sets



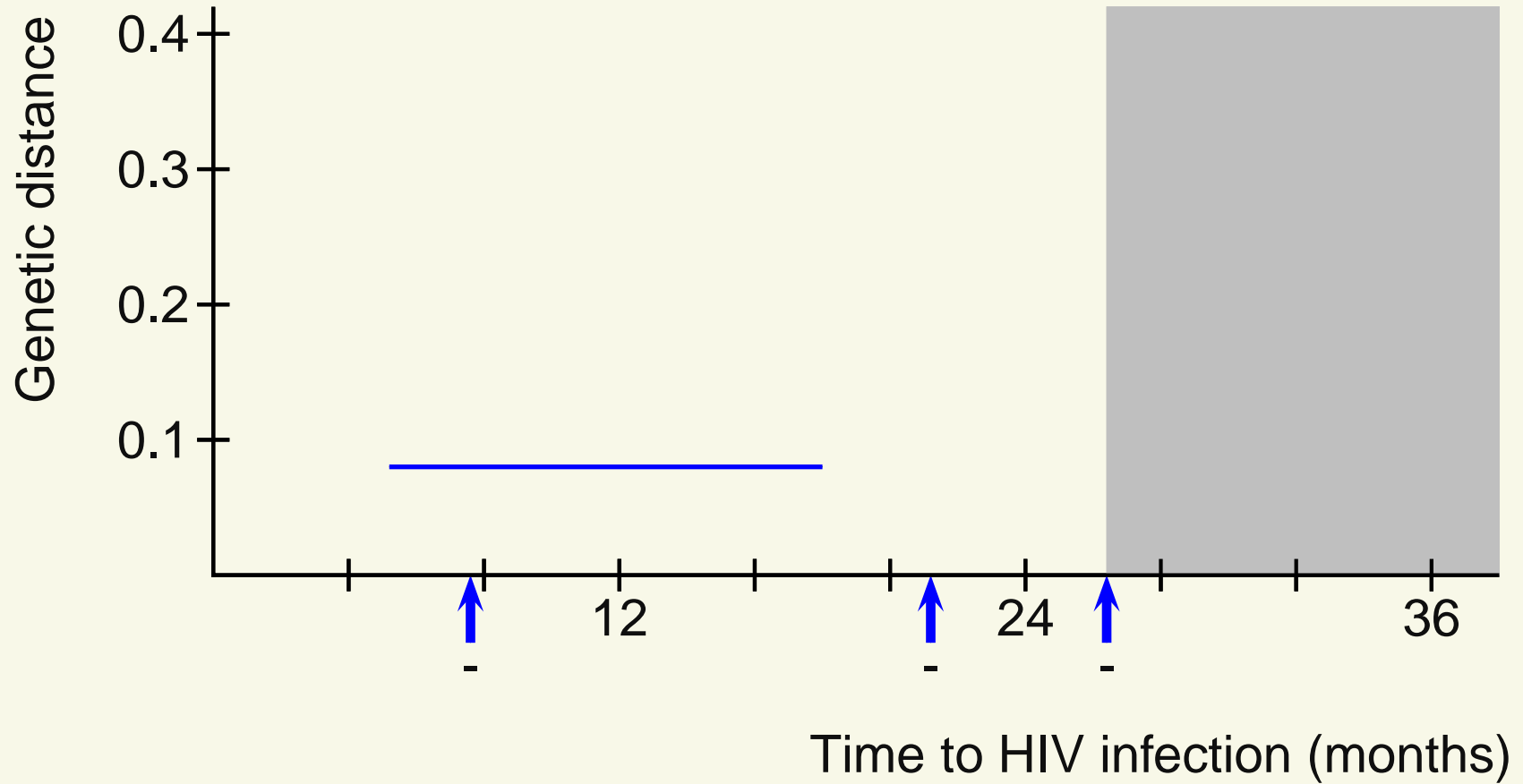
Observed sets



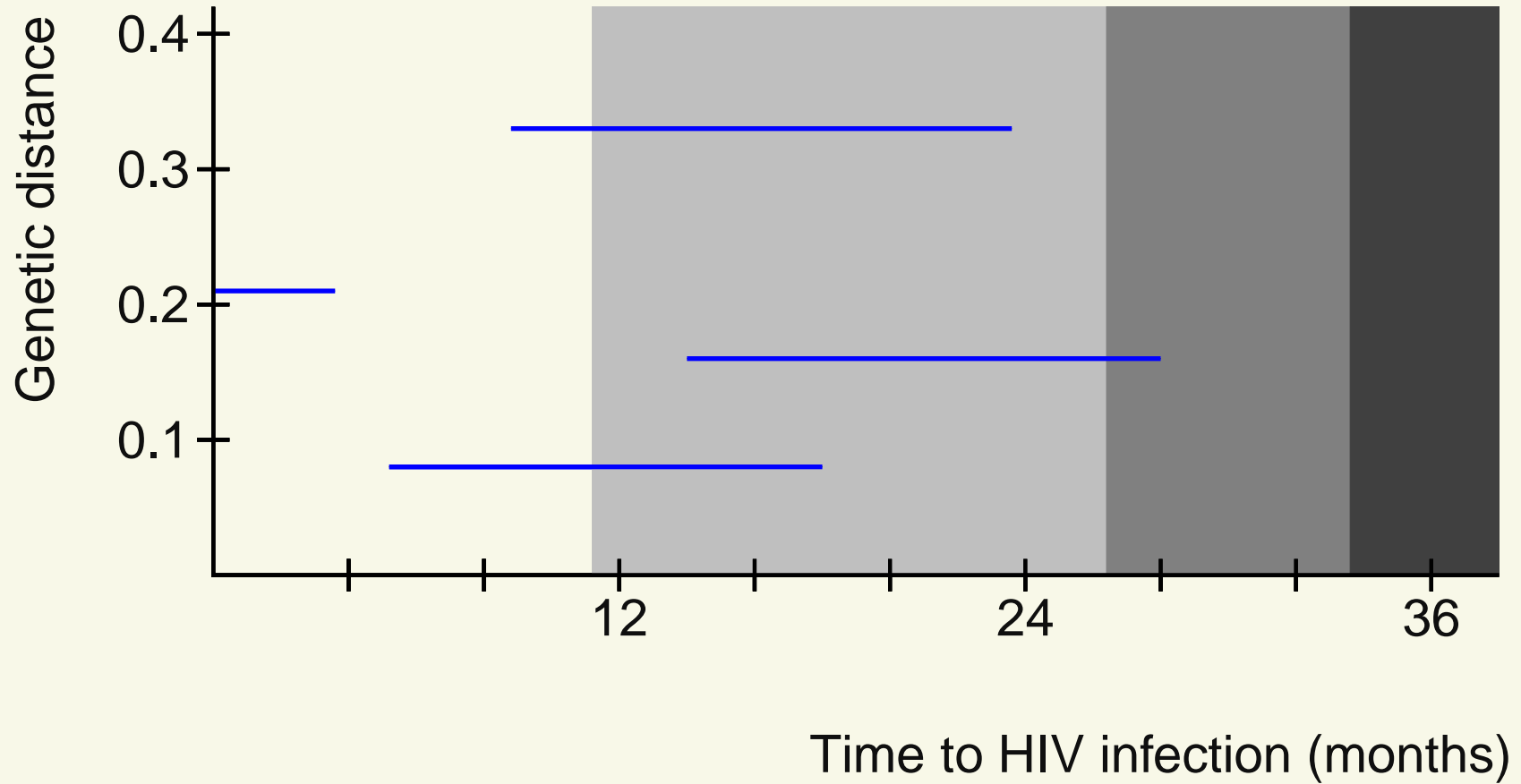
Observed sets



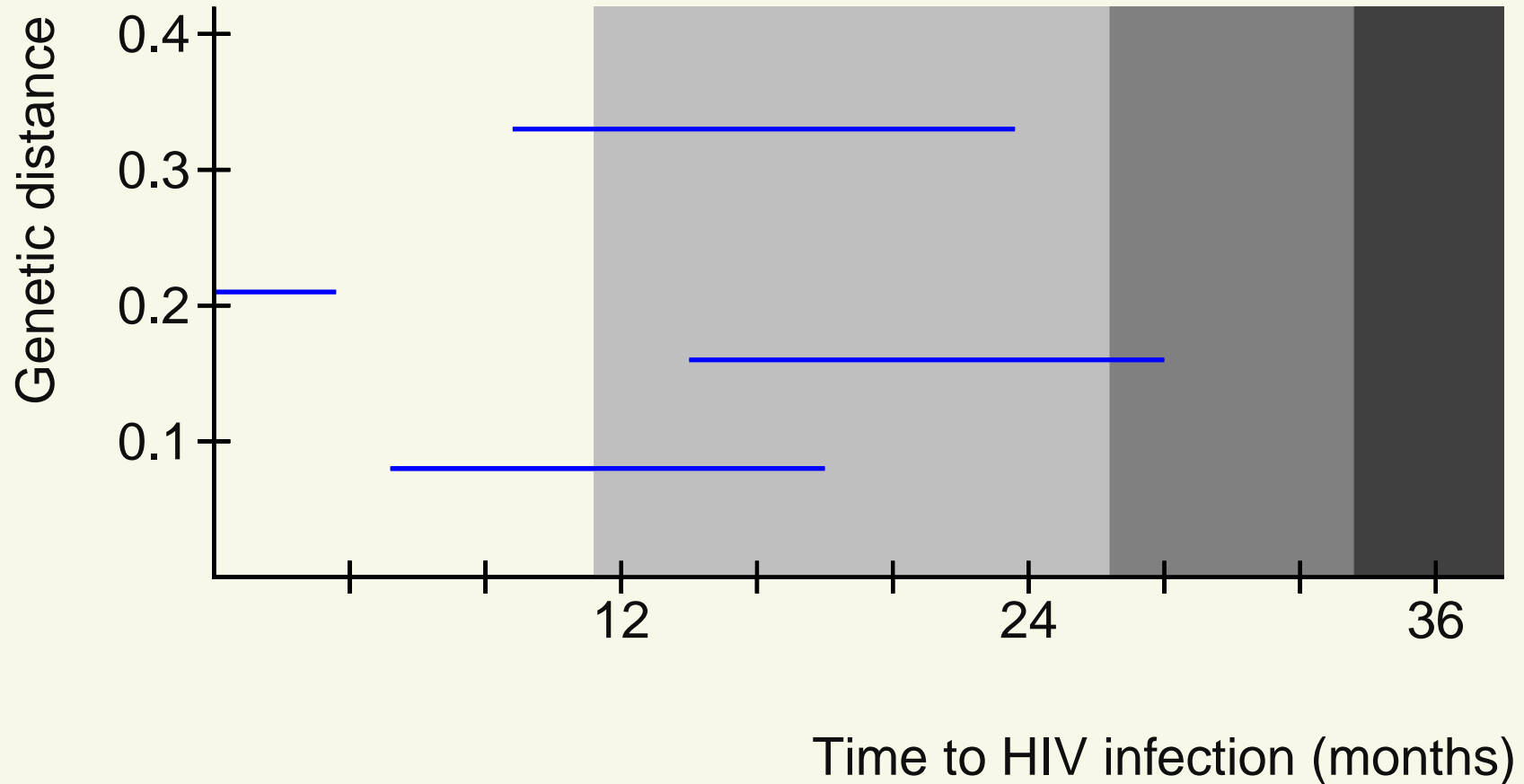
Observed sets



Observed sets



Observed sets



If Y is continuous, line segments do not overlap.

Model

- $(X, Y) \sim F_0, X \sim F_{0X}, Y \sim F_{0Y}$.
- We want to estimate F_0 .

Model

- $(X, Y) \sim F_0, X \sim F_{0X}, Y \sim F_{0Y}$.
- We want to estimate F_0 .
- X is interval censored. Instead of X , we observe $(\mathbf{T}, \mathbf{\Delta})$, where
 - $\mathbf{T} = (T_1, \dots, T_k), 0 < T_1 < \dots < T_k$ vector of observation times. The number k is fixed.
 - $\mathbf{\Delta} = (\Delta_1, \dots, \Delta_{k+1})$ and $\Delta_j = 1\{T_{j-1} < X \leq T_j\}$ for $j = 1, \dots, k + 1$ ($T_0 = 0, T_{k+1} = \infty$)
 - \mathbf{T} is independent of (X, Y) . $\mathbf{T} \sim G$.

Model

- $(X, Y) \sim F_0, X \sim F_{0X}, Y \sim F_{0Y}$.
- We want to estimate F_0 .
- X is interval censored. Instead of X , we observe $(\mathbf{T}, \mathbf{\Delta})$, where
 - $\mathbf{T} = (T_1, \dots, T_k), 0 < T_1 < \dots < T_k$ vector of observation times. The number k is fixed.
 - $\mathbf{\Delta} = (\Delta_1, \dots, \Delta_{k+1})$ and $\Delta_j = 1\{T_{j-1} < X \leq T_j\}$ for $j = 1, \dots, k + 1$ ($T_0 = 0, T_{k+1} = \infty$)
 - \mathbf{T} is independent of (X, Y) . $\mathbf{T} \sim G$.
- Y is a continuous mark variable:
 - It can only be observed if $X \leq T_k$, i.e., if $\Delta_+ = \sum_{j=1}^k \Delta_j = 1$.
 - Instead of Y , we observe $Z = \Delta_+ Y$.

Model

- $(X, Y) \sim F_0, X \sim F_{0X}, Y \sim F_{0Y}$.
- We want to estimate F_0 .
- X is interval censored. Instead of X , we observe $(\mathbf{T}, \mathbf{\Delta})$, where
 - $\mathbf{T} = (T_1, \dots, T_k), 0 < T_1 < \dots < T_k$ vector of observation times. The number k is fixed.
 - $\mathbf{\Delta} = (\Delta_1, \dots, \Delta_{k+1})$ and $\Delta_j = 1\{T_{j-1} < X \leq T_j\}$ for $j = 1, \dots, k+1$ ($T_0 = 0, T_{k+1} = \infty$)
 - \mathbf{T} is independent of (X, Y) . $\mathbf{T} \sim G$.
- Y is a continuous mark variable:
 - It can only be observed if $X \leq T_k$, i.e., if $\Delta_+ = \sum_{j=1}^k \Delta_j = 1$.
 - Instead of Y , we observe $Z = \Delta_+ Y$.
- Observed data: $W = (\mathbf{T}, \mathbf{\Delta}, Z)$.

Model

- $(X, Y) \sim F_0, X \sim F_{0X}, Y \sim F_{0Y}$.
- We want to estimate F_0 .
- X is interval censored. Instead of X , we observe $(\mathbf{T}, \mathbf{\Delta})$, where
 - $\mathbf{T} = (T_1, \dots, T_k), 0 < T_1 < \dots < T_k$ vector of observation times. The number k is fixed.
 - $\mathbf{\Delta} = (\Delta_1, \dots, \Delta_{k+1})$ and $\Delta_j = 1\{T_{j-1} < X \leq T_j\}$ for $j = 1, \dots, k + 1$ ($T_0 = 0, T_{k+1} = \infty$)
 - \mathbf{T} is independent of (X, Y) . $\mathbf{T} \sim G$.
- Y is a continuous mark variable:
 - It can only be observed if $X \leq T_k$, i.e., if $\Delta_+ = \sum_{j=1}^k \Delta_j = 1$.
 - Instead of Y , we observe $Z = \Delta_+ Y$.
- Observed data: $W = (\mathbf{T}, \mathbf{\Delta}, Z)$.
- We study the nonparametric maximum likelihood estimator (MLE) for F_0 , based on n i.i.d. observations W_1, \dots, W_n .

Outline

- MLE is inconsistent in general
- Simple method to repair inconsistency
- Simulation results of the MLE and the repaired MLE

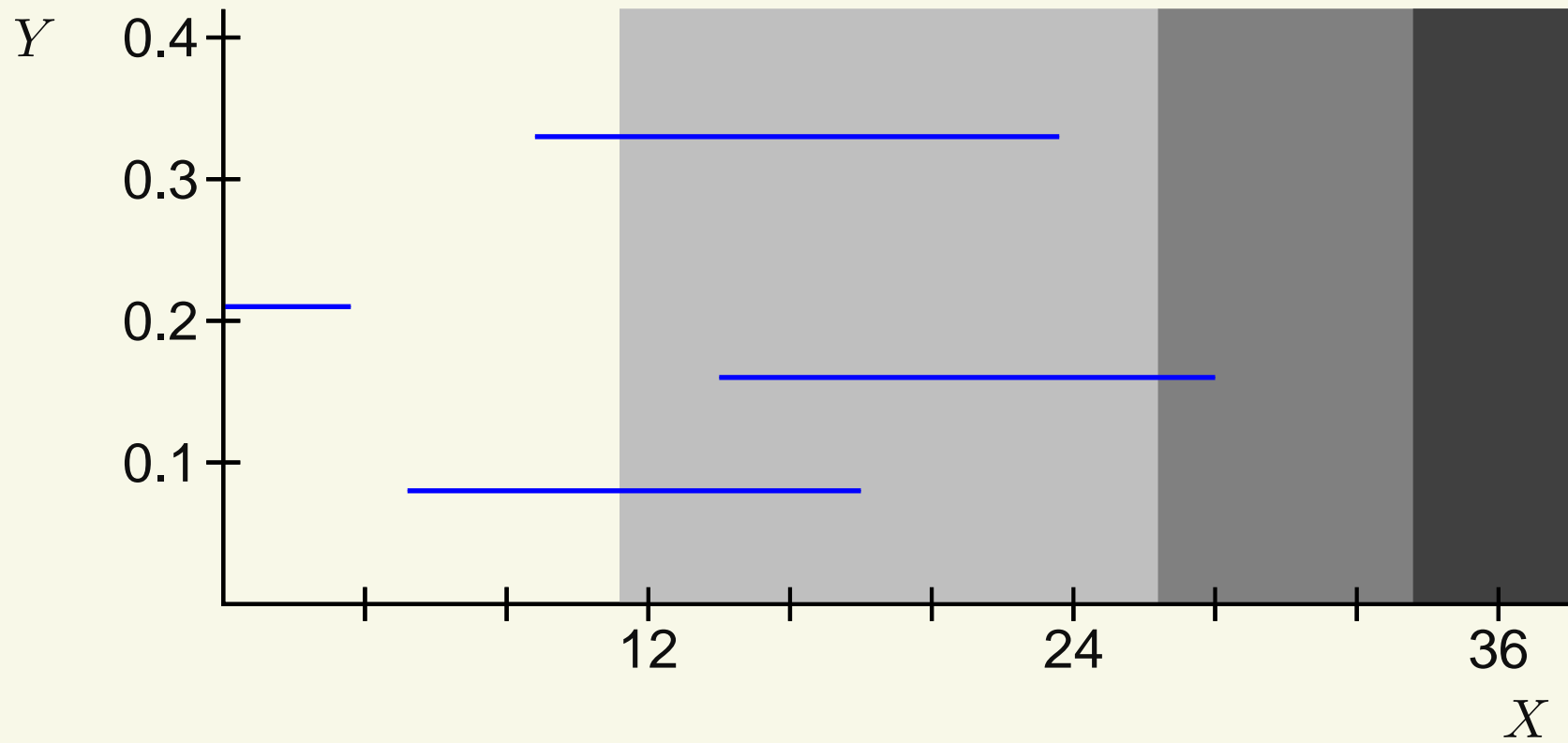
Inconsistency

- We suspected inconsistency because:
 - The observed sets in our model can take the form of line segments.
 - This was also the case in the models with inconsistent MLEs studied by Van der Laan (1996) and Maathuis (2003).

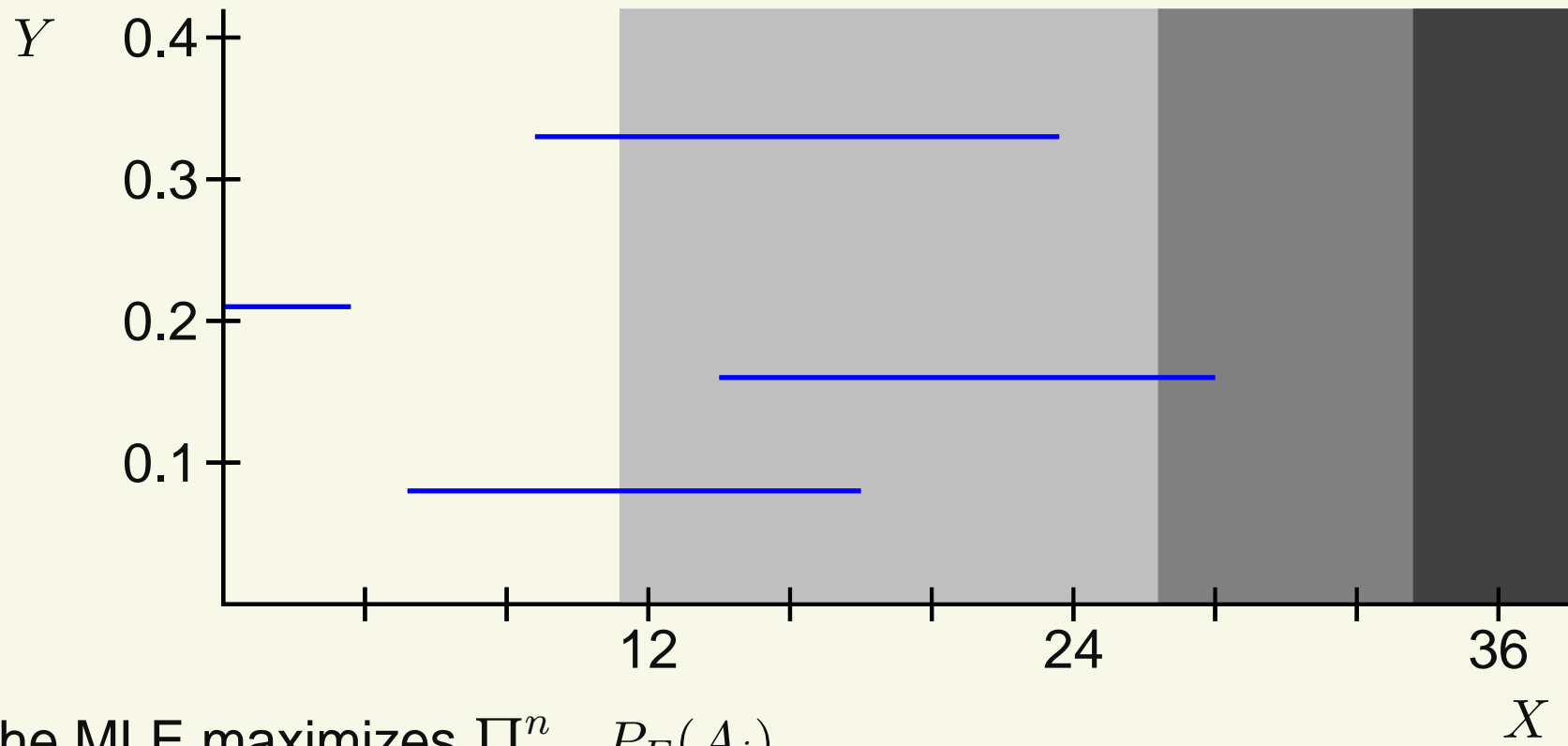
Inconsistency

- We suspected inconsistency because:
 - The observed sets in our model can take the form of line segments.
 - This was also the case in the models with inconsistent MLEs studied by Van der Laan (1996) and Maathuis (2003).
- How to prove inconsistency?
 - Derive explicit formula for MLE.
 - Express this formula in terms of empirical processes.
 - Then it is straightforward to derive the limit.
 - This leads to necessary and sufficient conditions for consistency.
 - These conditions force a relation between G and F_0 .
 - Such a relation does typically not hold.

The MLE



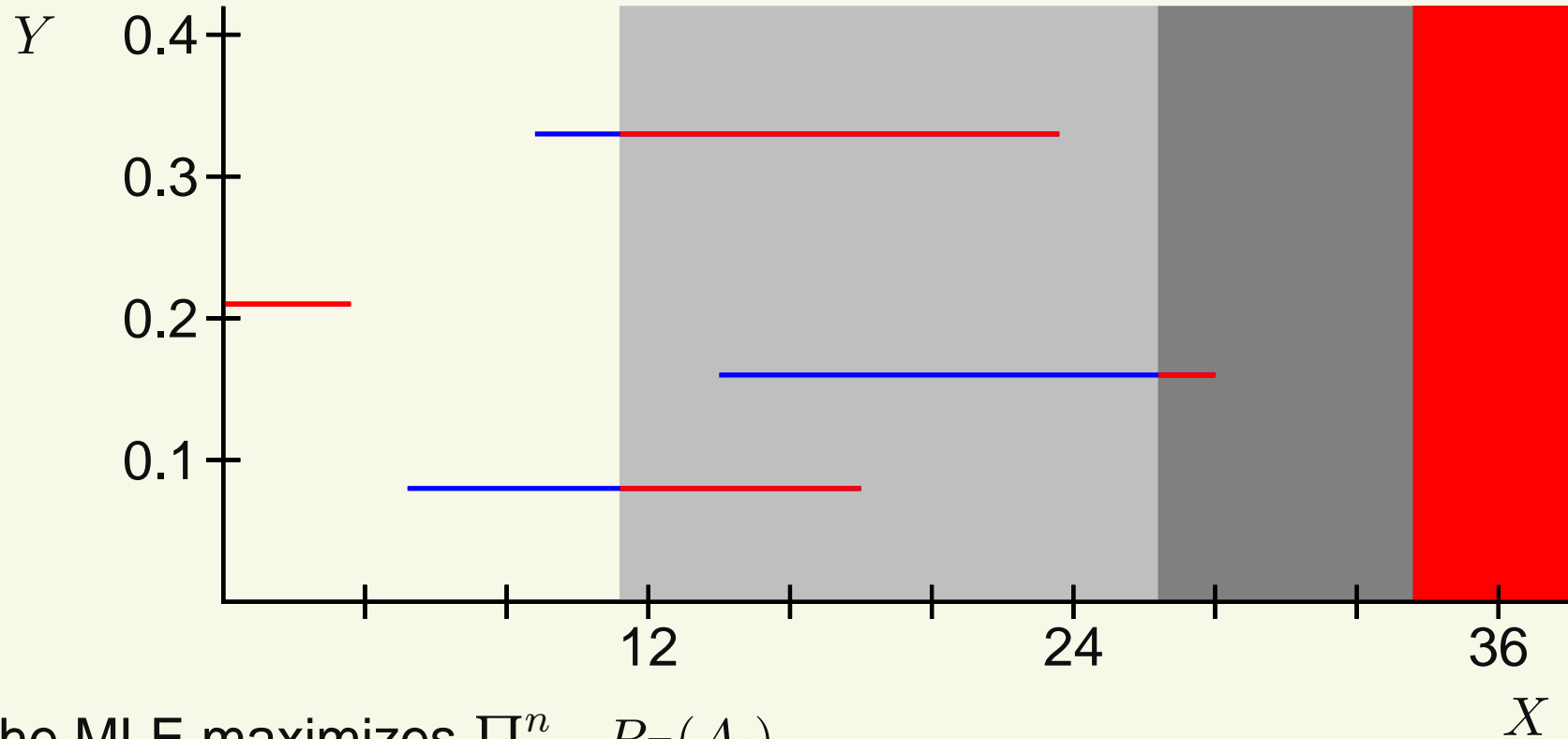
The MLE



The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

X

The MLE

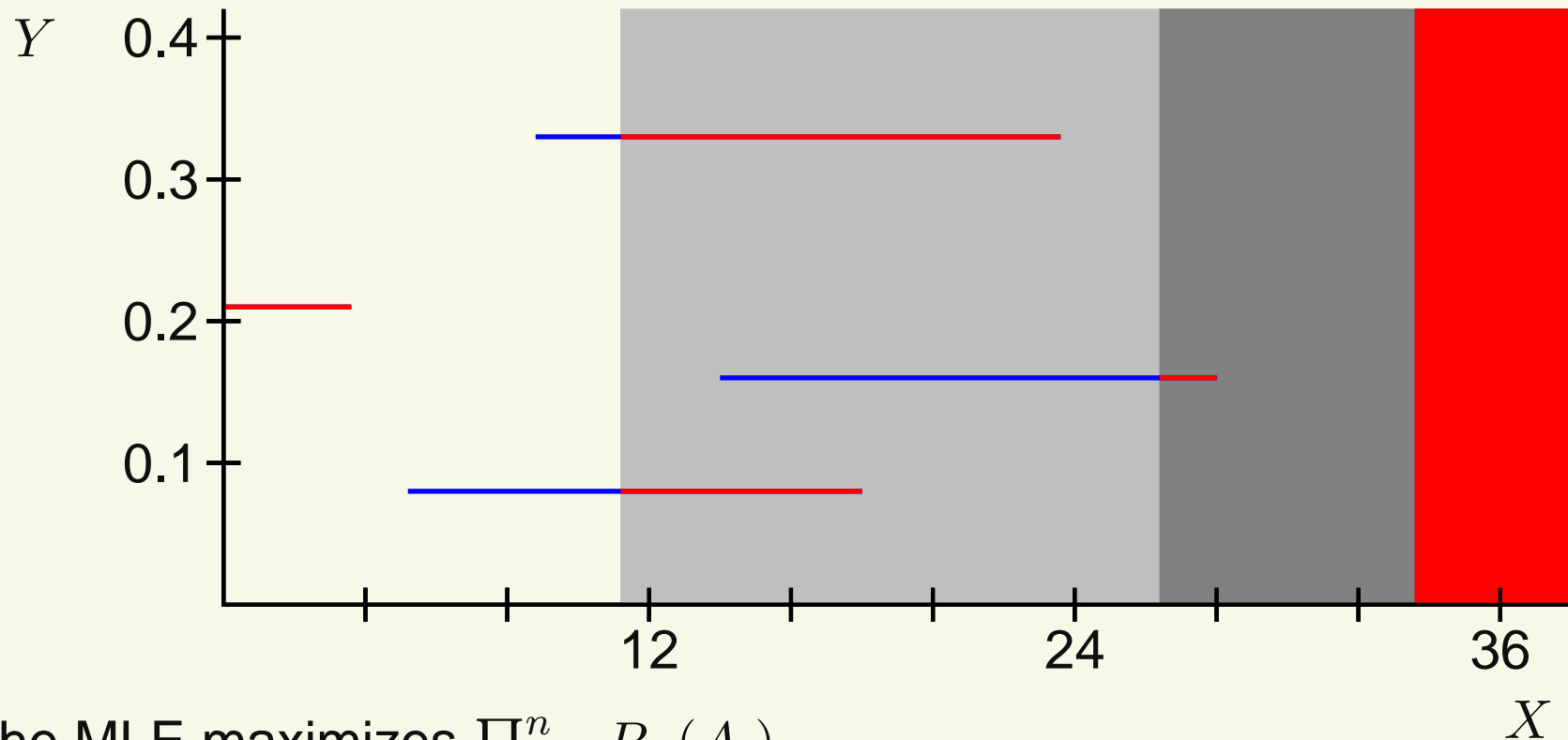


The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

X

The MLE

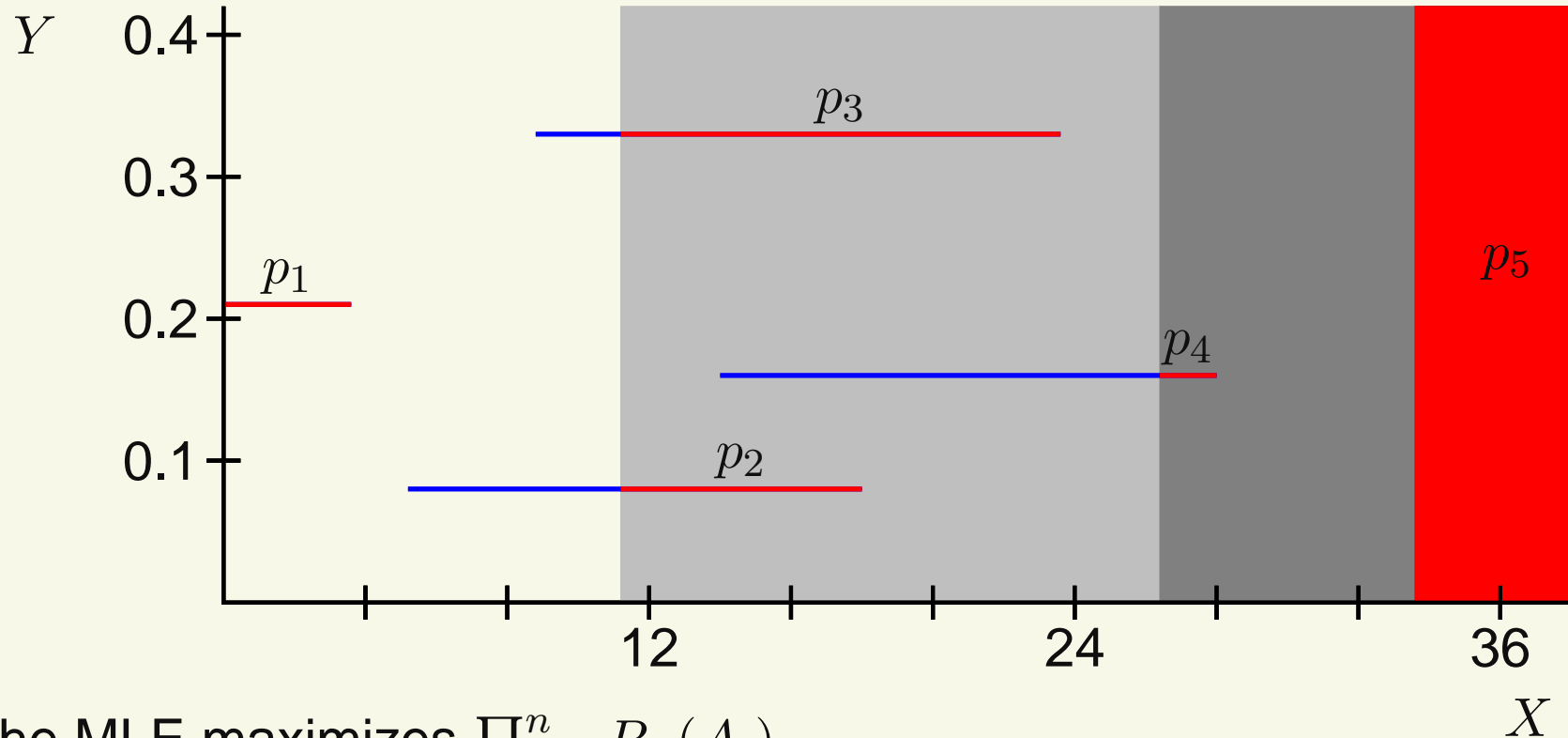


The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

The MLE is non-unique.

The MLE

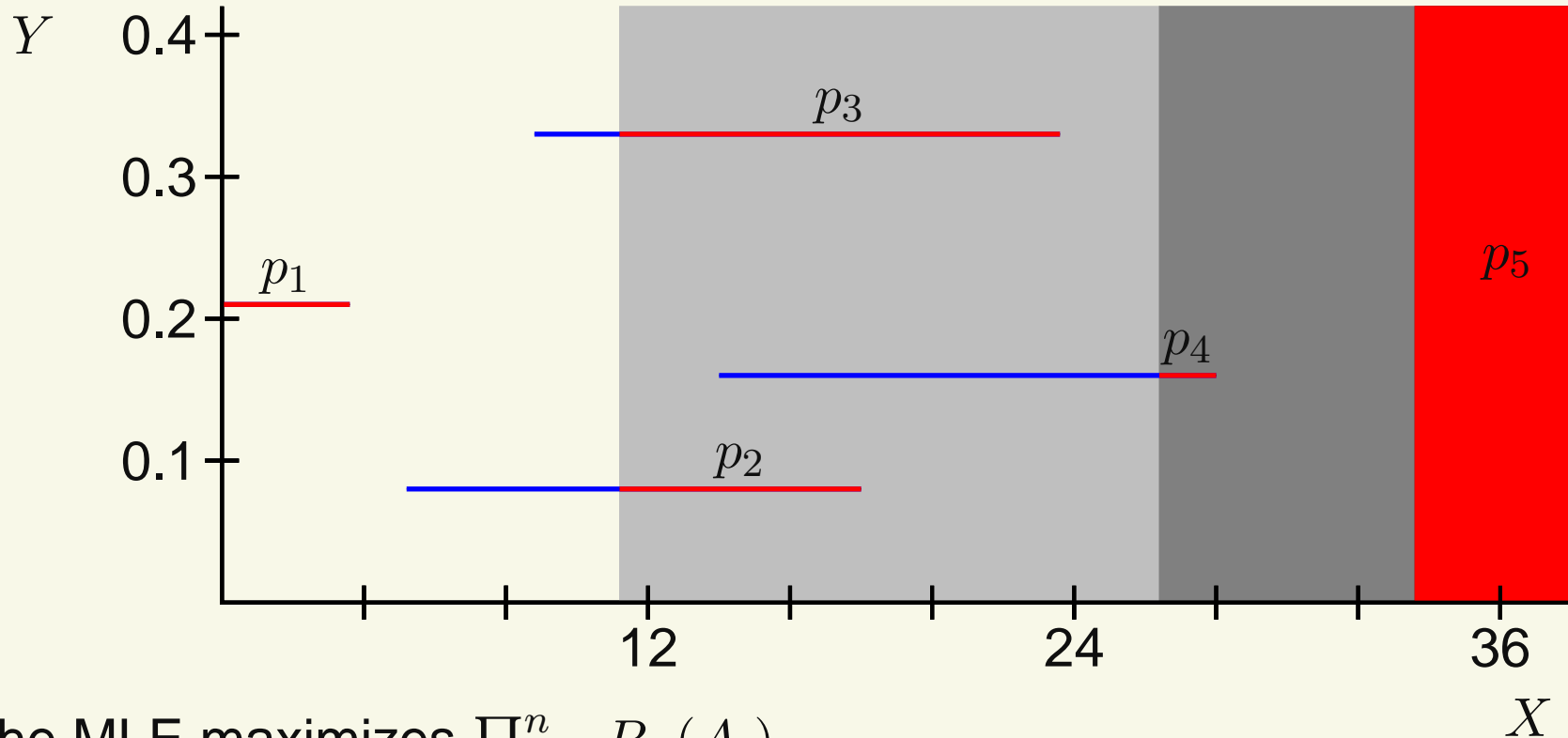


The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

The MLE is non-unique.

The MLE



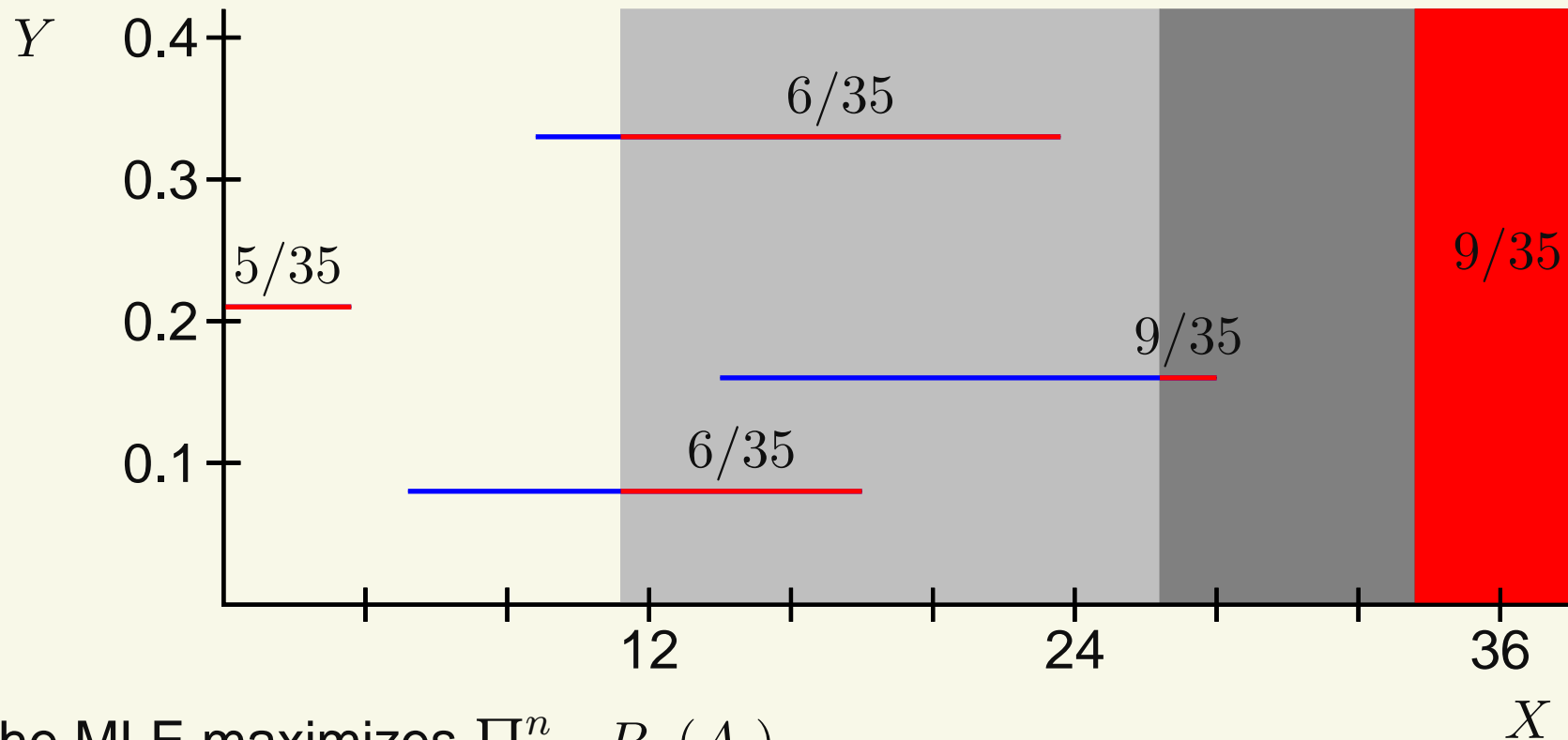
The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

The MLE is non-unique.

$$\prod_{i=1}^n P_F(A_i) = p_1 p_2 p_3 p_4 (p_2 + p_3 + p_4 + p_5) (p_4 + p_5) p_5.$$

The MLE



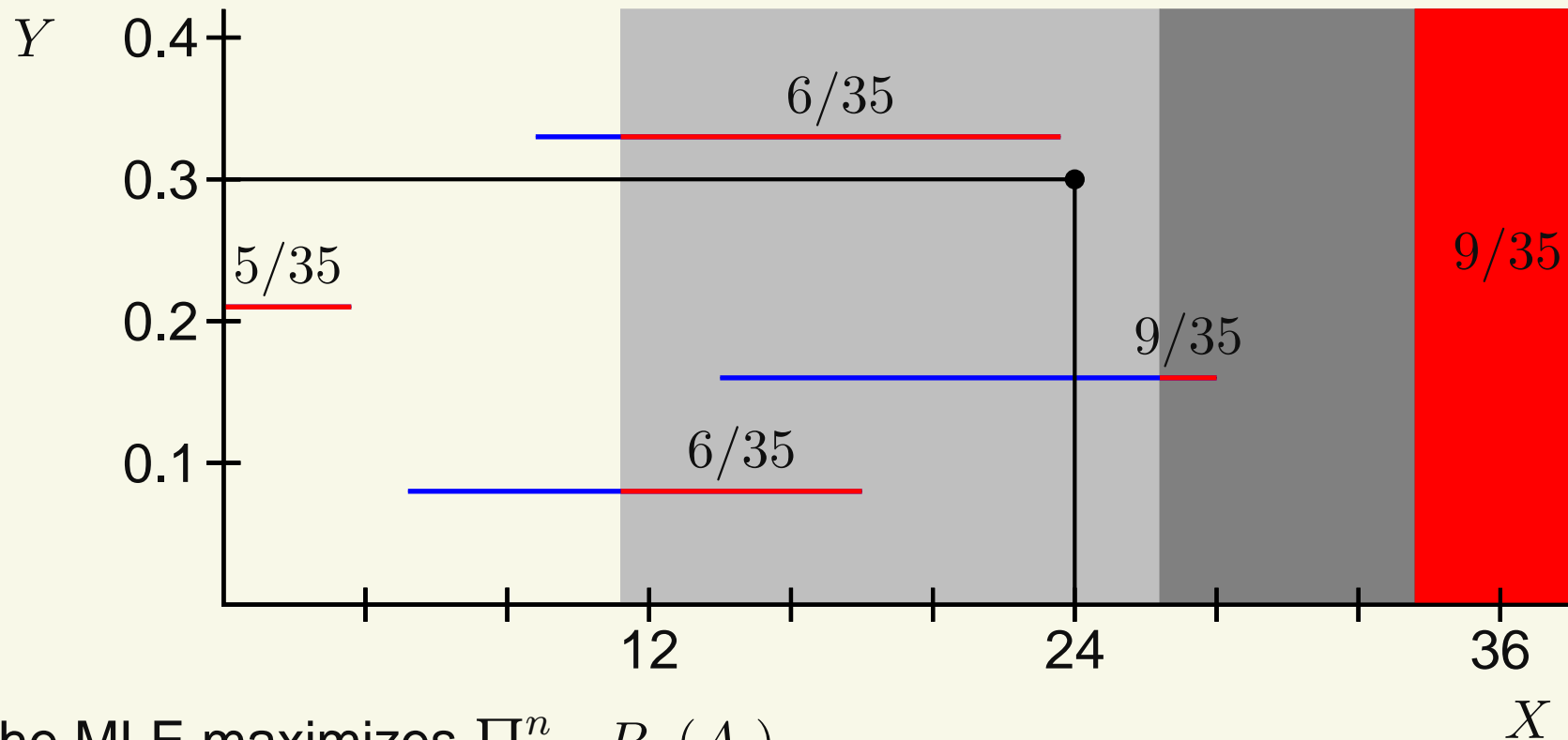
The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

The MLE is non-unique.

$$\prod_{i=1}^n P_F(A_i) = p_1 p_2 p_3 p_4 (p_2 + p_3 + p_4 + p_5) (p_4 + p_5) p_5.$$

The MLE



The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

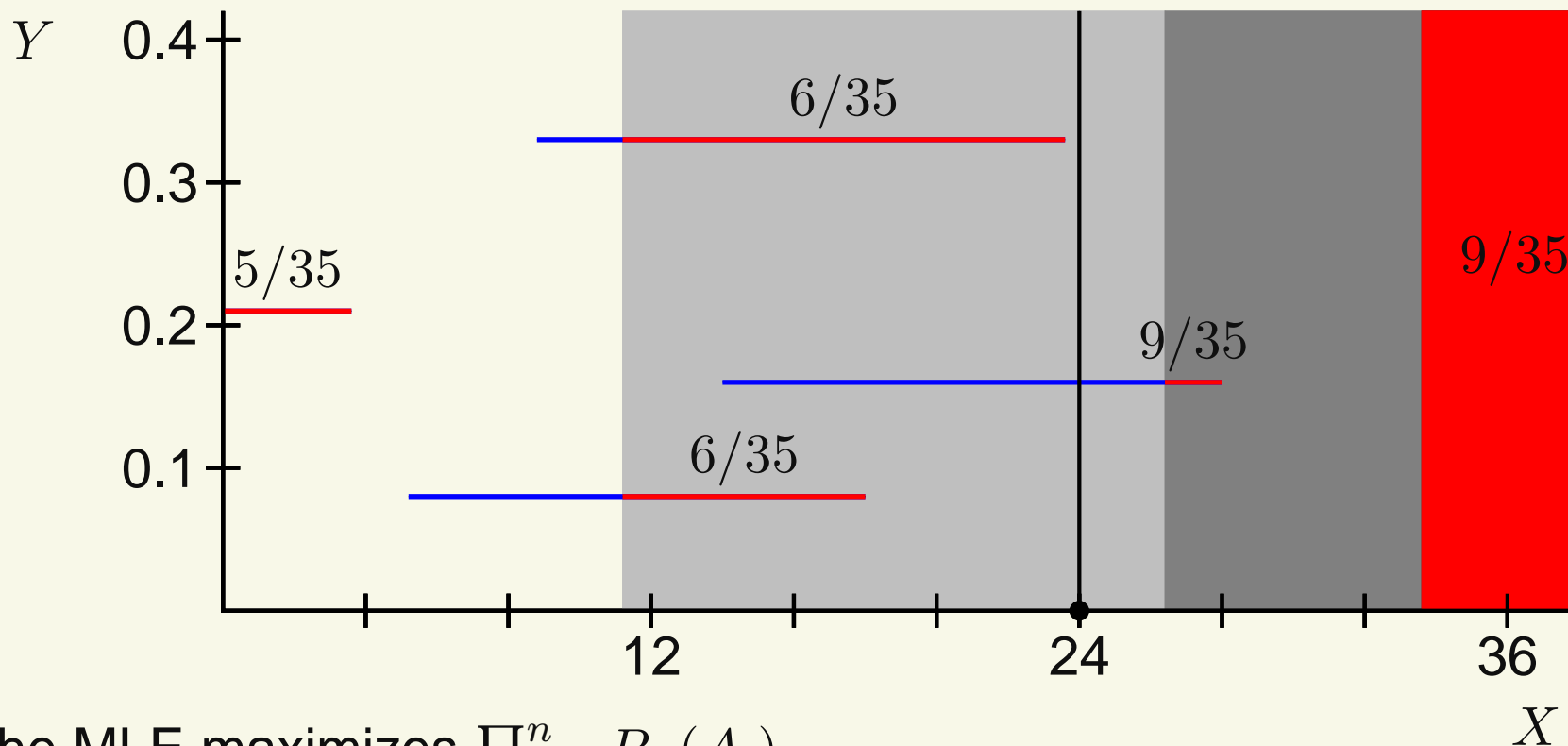
The MLE can only put mass in the maximal intersections.

The MLE is non-unique.

$$\prod_{i=1}^n P_F(A_i) = p_1 p_2 p_3 p_4 (p_2 + p_3 + p_4 + p_5) (p_4 + p_5) p_5.$$

Obtain \hat{F}_n and \hat{F}_{X_n} by summing up mass.

The MLE



The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

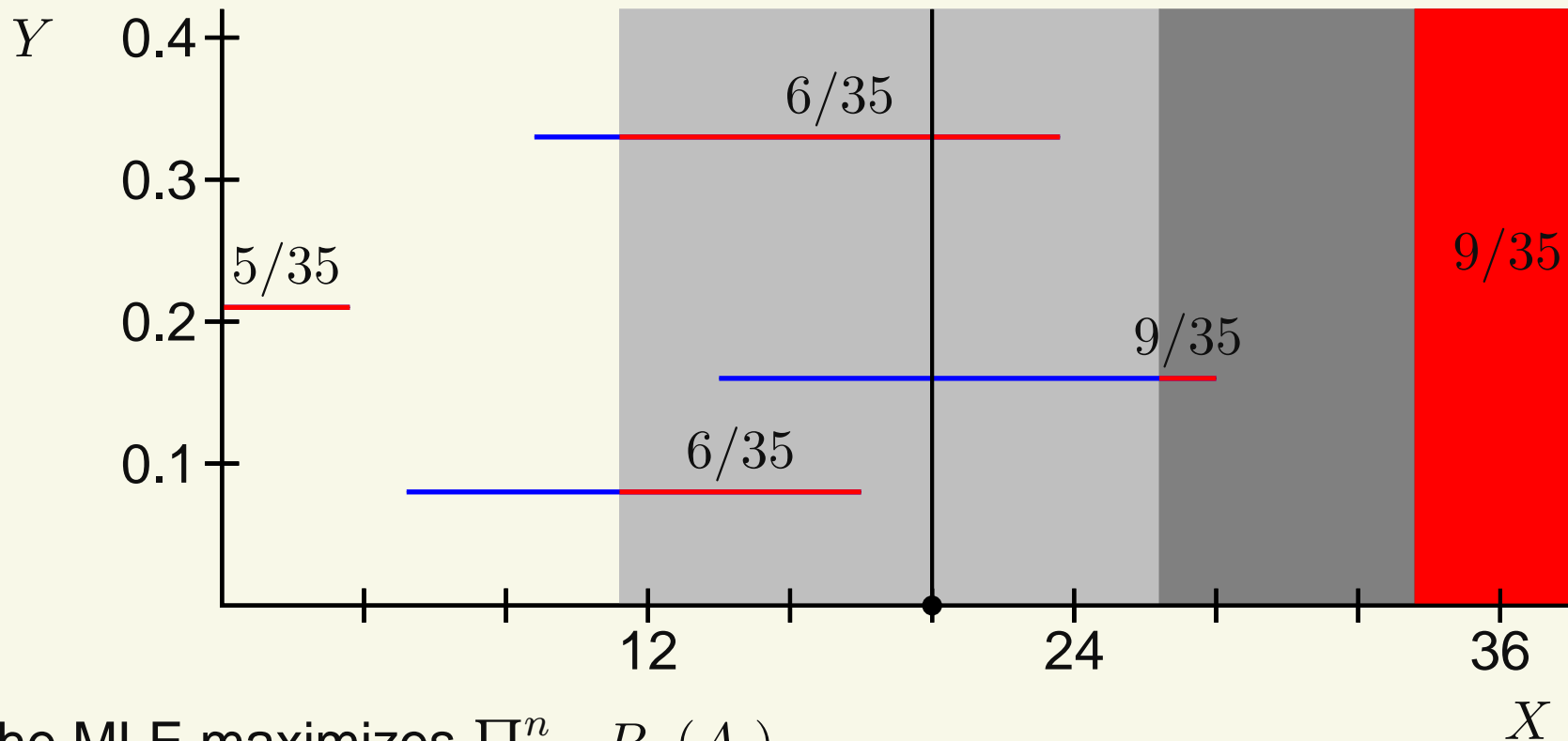
The MLE can only put mass in the maximal intersections.

The MLE is non-unique.

$$\prod_{i=1}^n P_F(A_i) = p_1 p_2 p_3 p_4 (p_2 + p_3 + p_4 + p_5) (p_4 + p_5) p_5.$$

Obtain \hat{F}_n and \hat{F}_{X_n} by summing up mass.

The MLE



The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

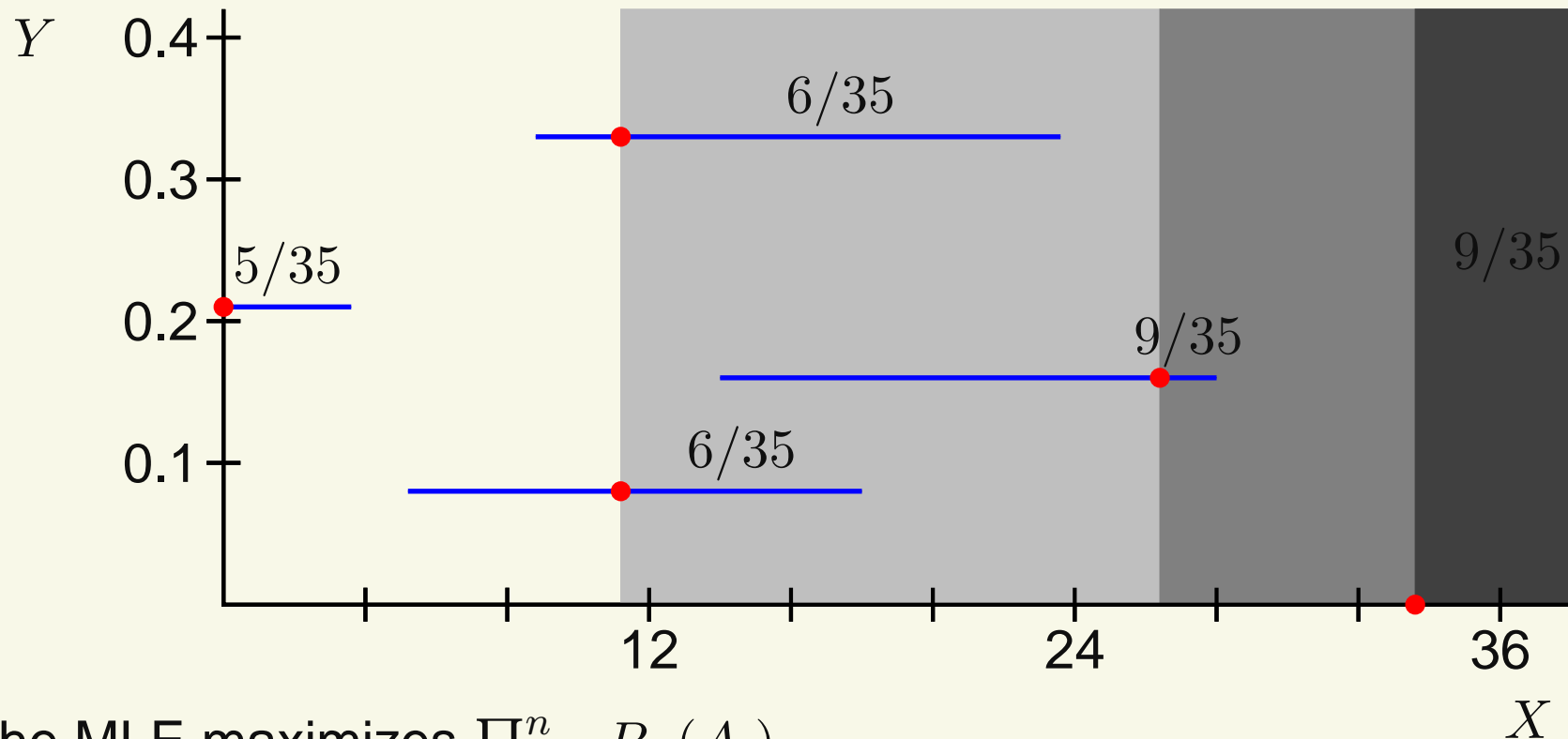
The MLE can only put mass in the maximal intersections.

The MLE is non-unique.

$$\prod_{i=1}^n P_F(A_i) = p_1 p_2 p_3 p_4 (p_2 + p_3 + p_4 + p_5) (p_4 + p_5) p_5.$$

Obtain \hat{F}_n and \hat{F}_{X_n} by summing up mass.

The MLE



The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

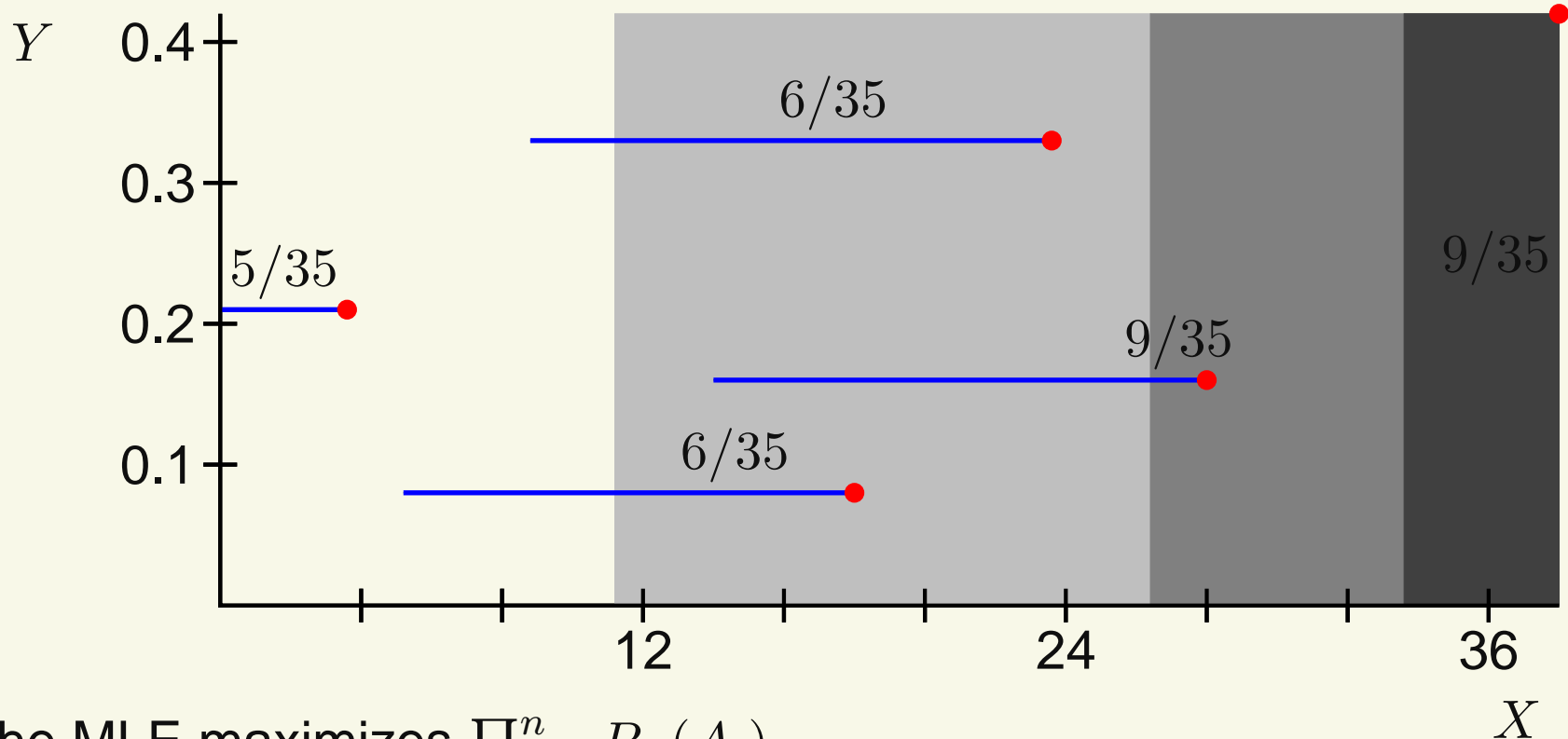
The MLE is non-unique.

$$\prod_{i=1}^n P_F(A_i) = p_1 p_2 p_3 p_4 (p_2 + p_3 + p_4 + p_5) (p_4 + p_5) p_5.$$

Obtain \hat{F}_n and \hat{F}_{X_n} by summing up mass.

Upper bounds $\hat{F}_n^u, \hat{F}_{X_n}^u$.

The MLE



The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

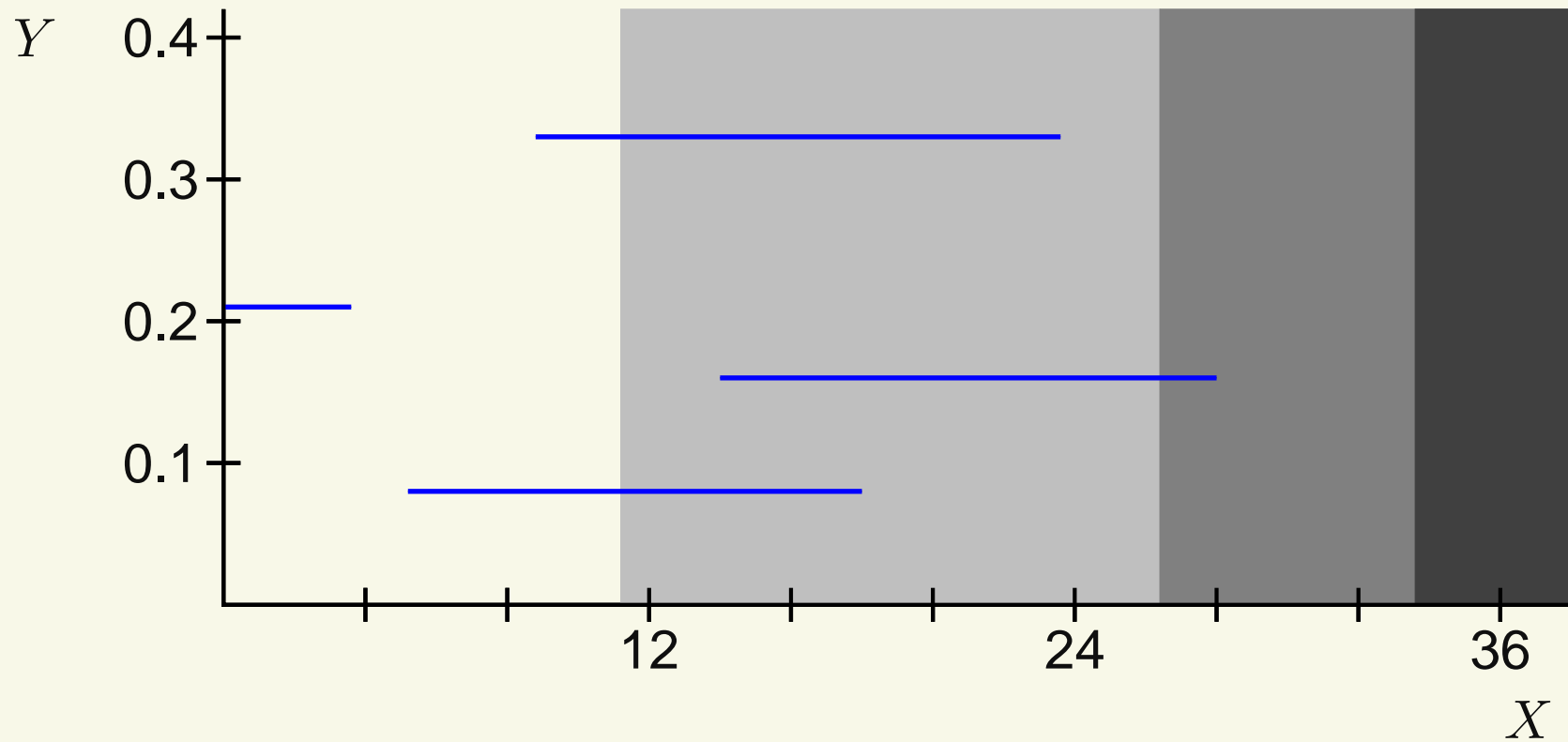
The MLE is non-unique.

$$\prod_{i=1}^n P_F(A_i) = p_1 p_2 p_3 p_4 (p_2 + p_3 + p_4 + p_5) (p_4 + p_5) p_5.$$

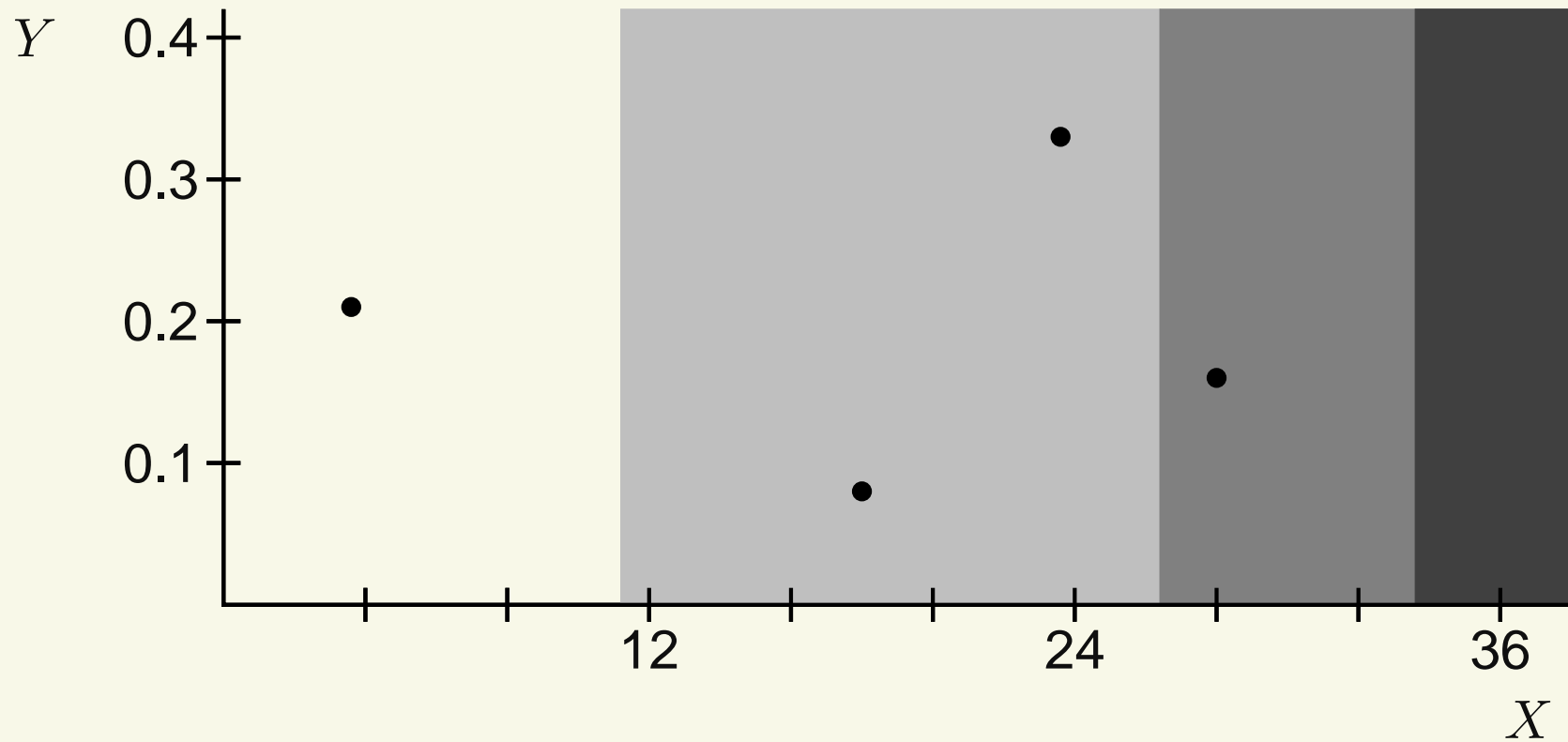
Obtain \hat{F}_n and \hat{F}_{X_n} by summing up mass.

Lower bounds $\hat{F}_n^\ell, \hat{F}_{X_n}^\ell$.

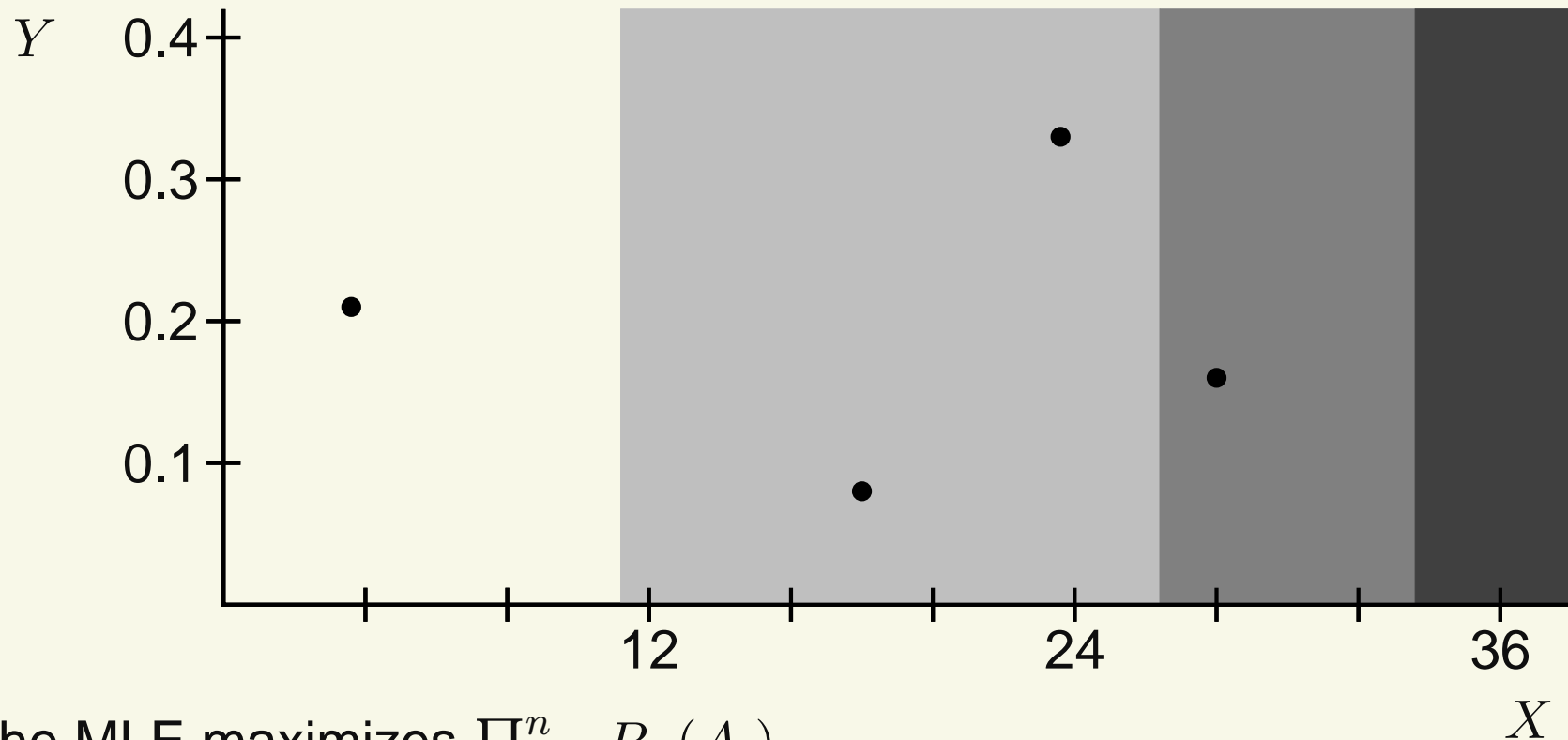
The MLE can be viewed as a right endpoint imputation estimator



The MLE can be viewed as a right endpoint imputation estimator



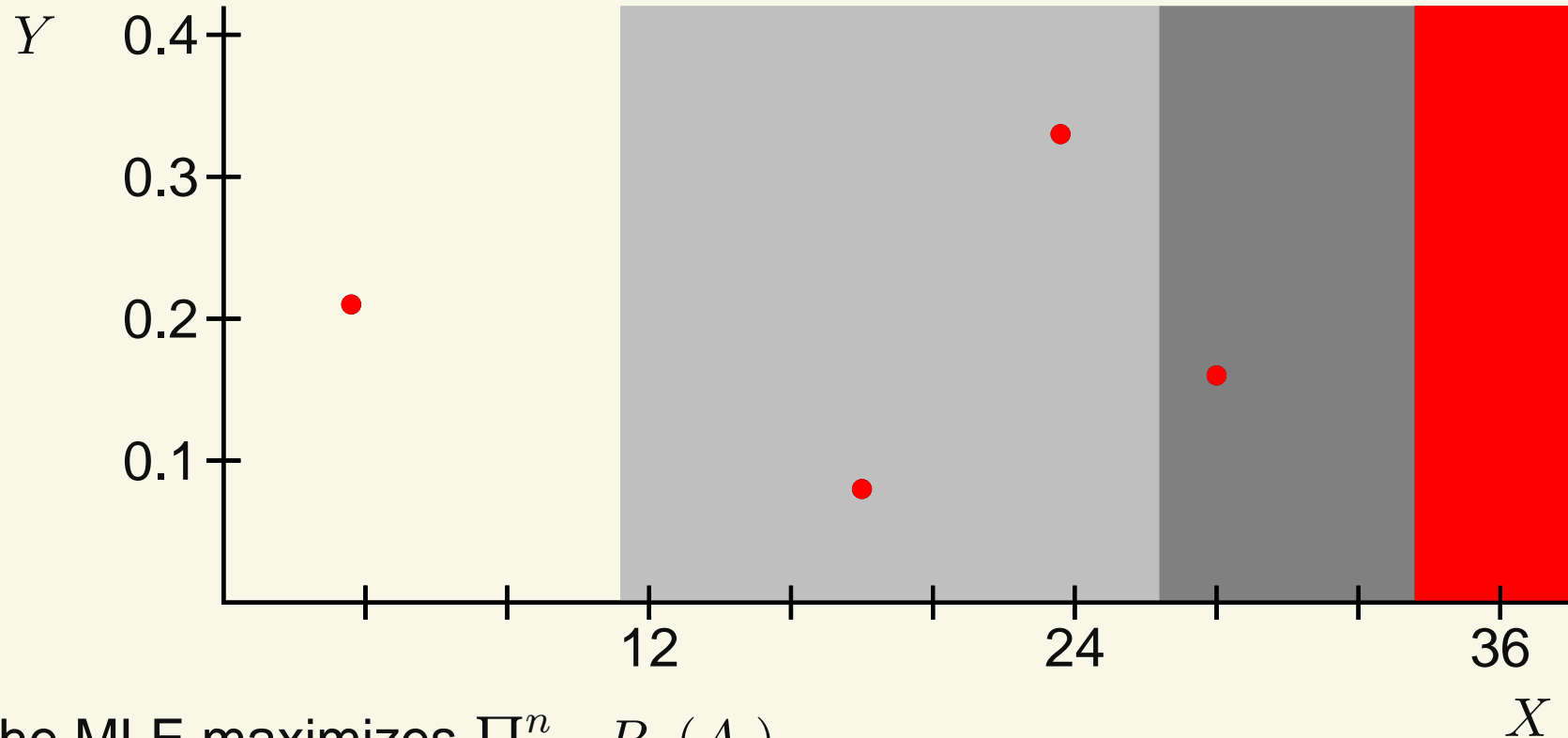
The MLE can be viewed as a right endpoint imputation estimator



The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

X

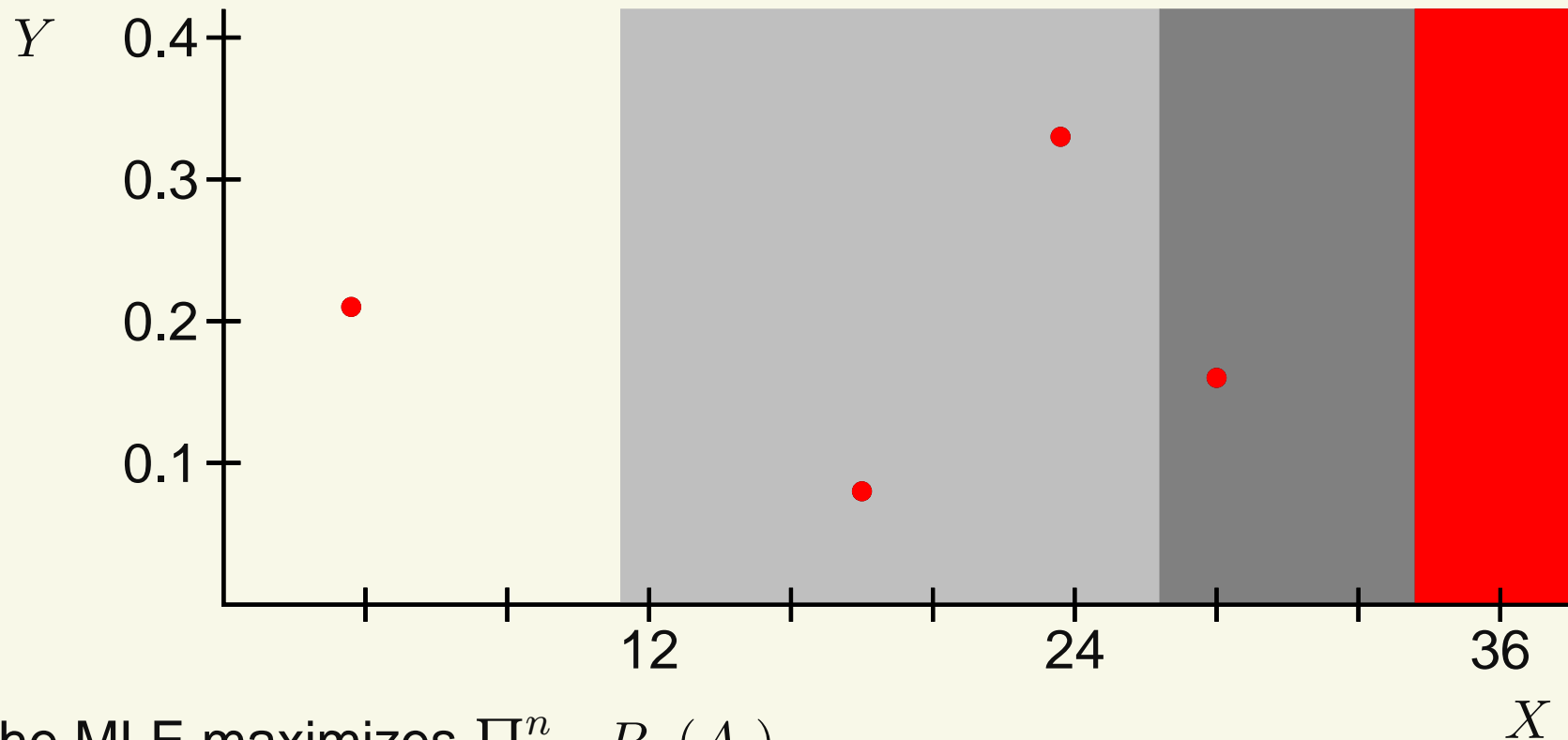
The MLE can be viewed as a right endpoint imputation estimator



The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

The MLE can be viewed as a right endpoint imputation estimator

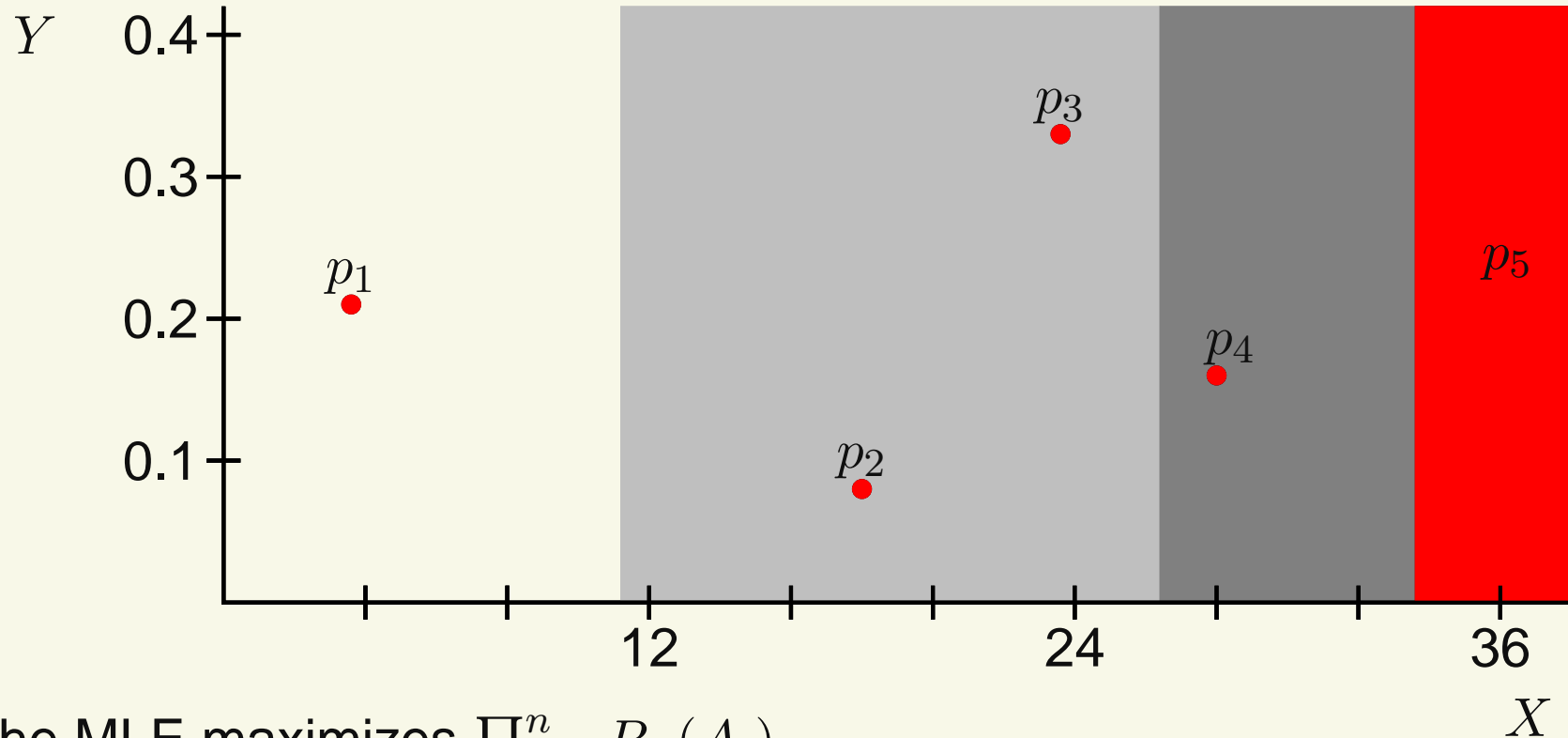


The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

The MLE may be non-unique.

The MLE can be viewed as a right endpoint imputation estimator

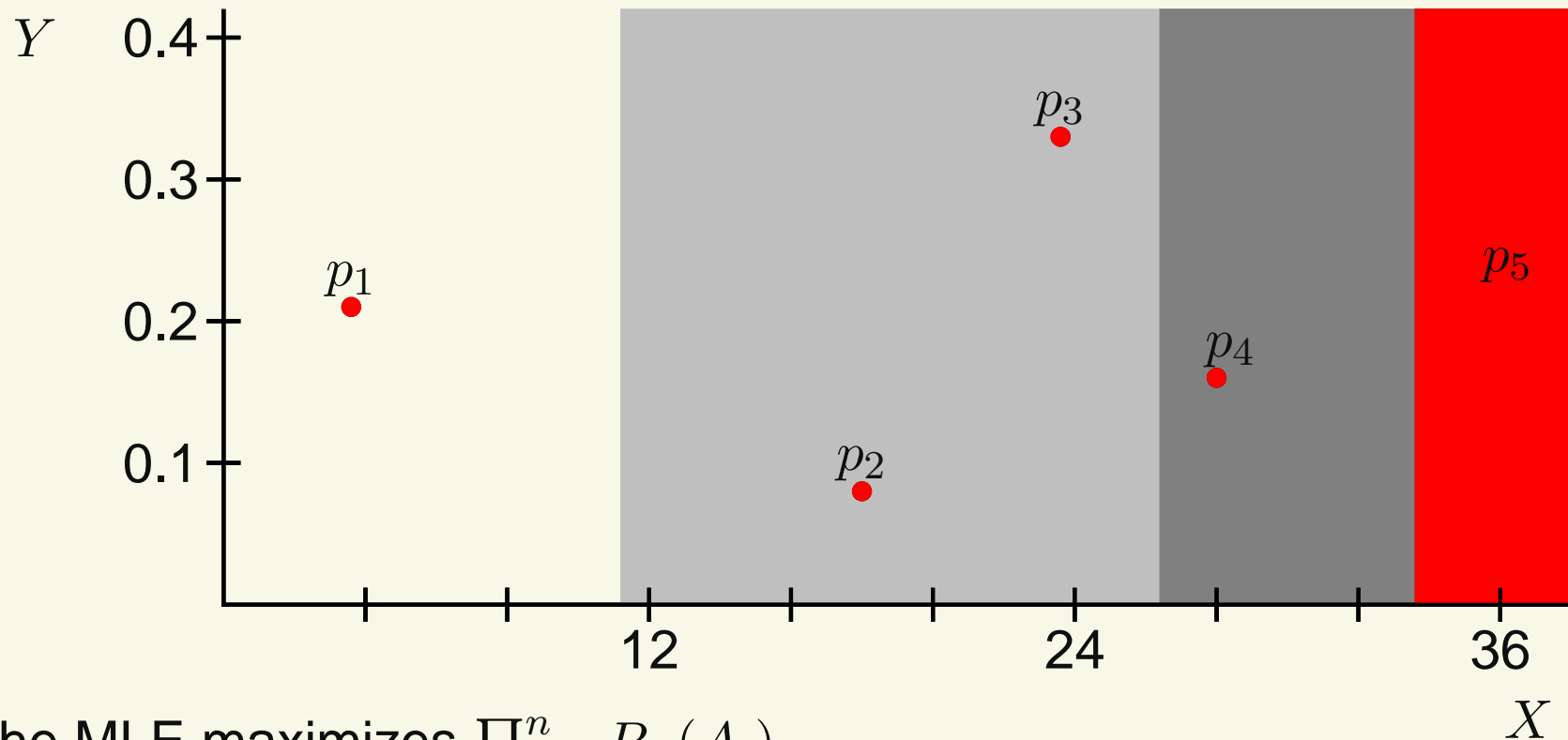


The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

The MLE may be non-unique.

The MLE can be viewed as a right endpoint imputation estimator



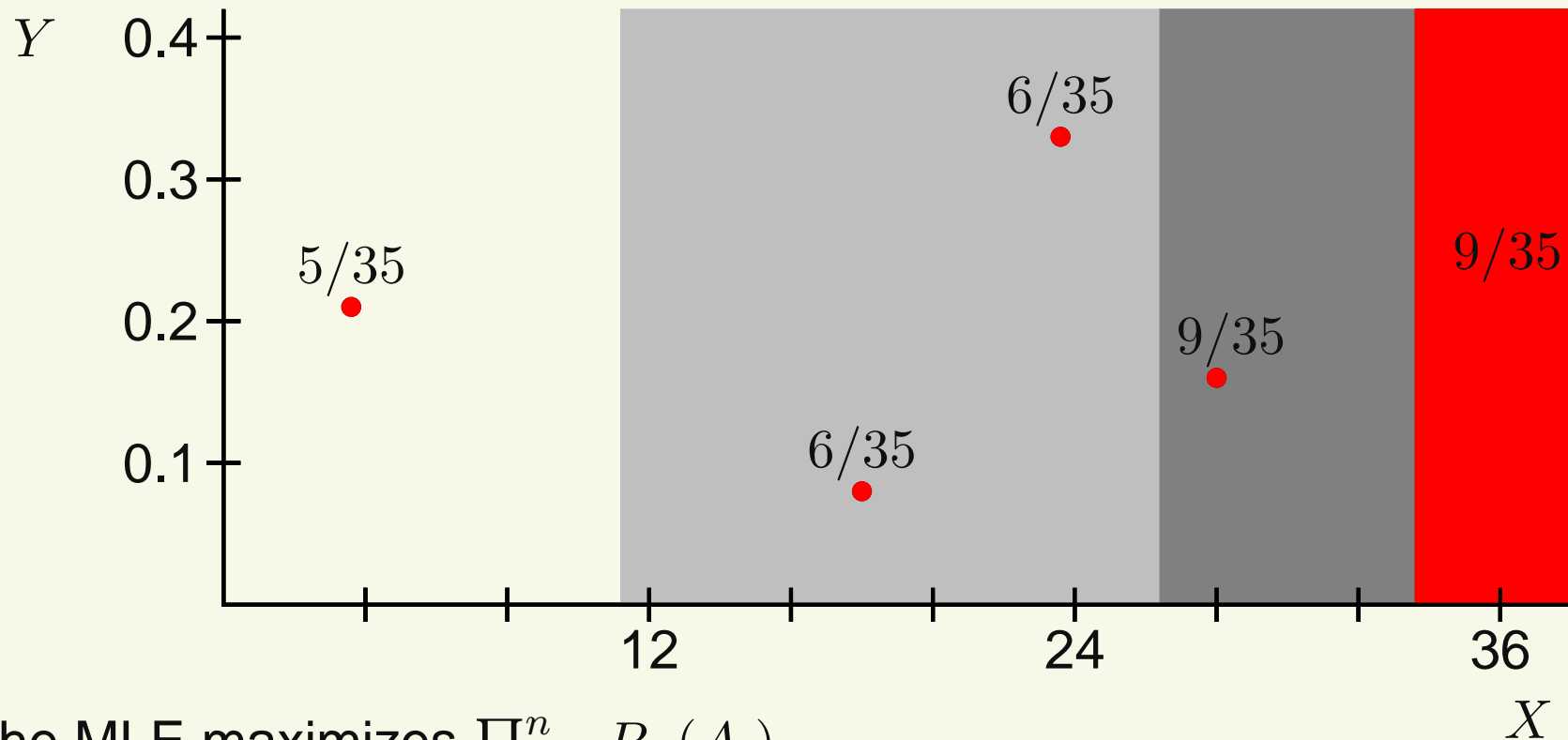
The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

The MLE may be non-unique.

$$\prod_{i=1}^n P_F(A_i) = p_1 p_2 p_3 p_4 (p_2 + p_3 + p_4 + p_5) (p_4 + p_5) p_5.$$

The MLE can be viewed as a right endpoint imputation estimator



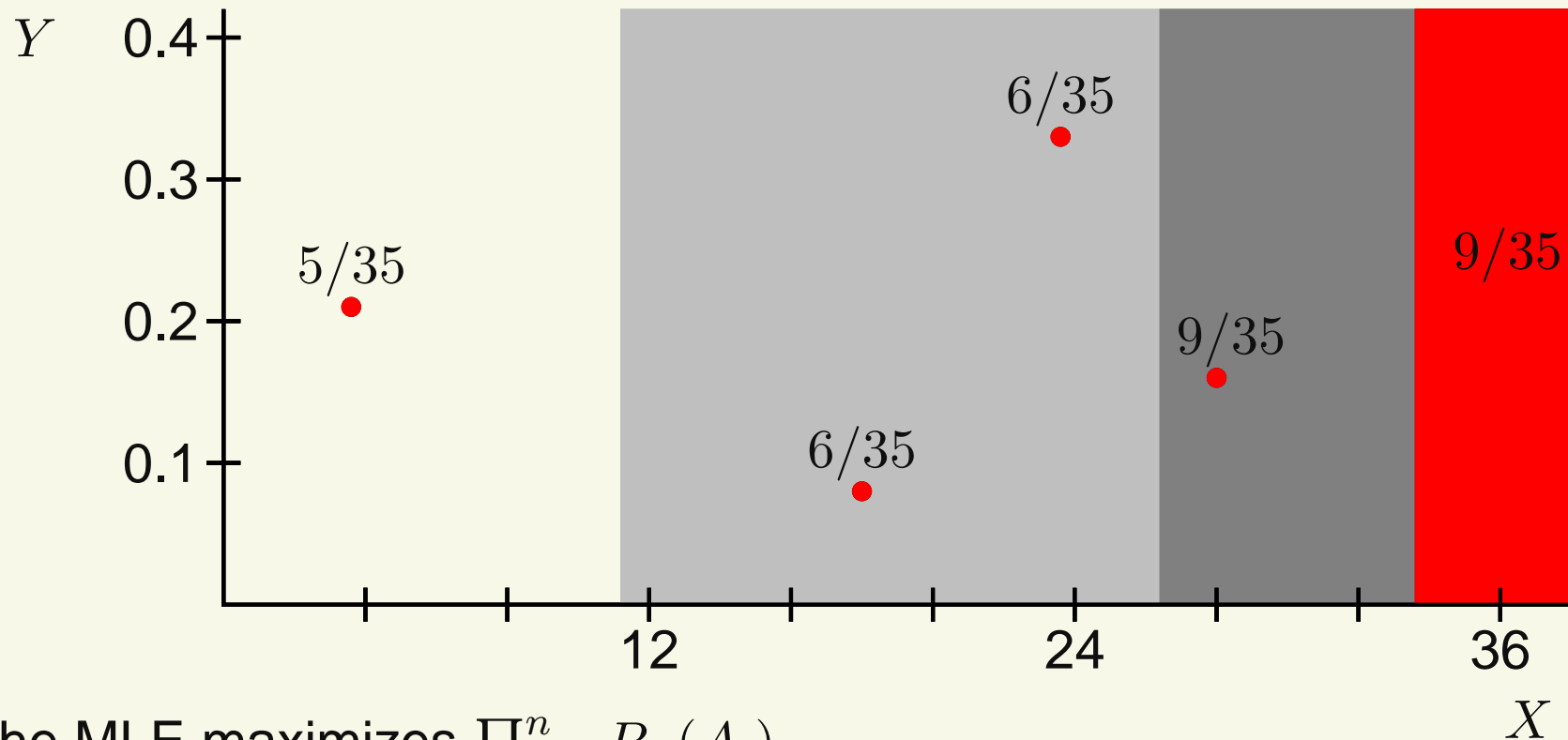
The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

The MLE may be non-unique.

$$\prod_{i=1}^n P_F(A_i) = p_1 p_2 p_3 p_4 (p_2 + p_3 + p_4 + p_5) (p_4 + p_5) p_5.$$

The MLE can be viewed as a right endpoint imputation estimator



The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

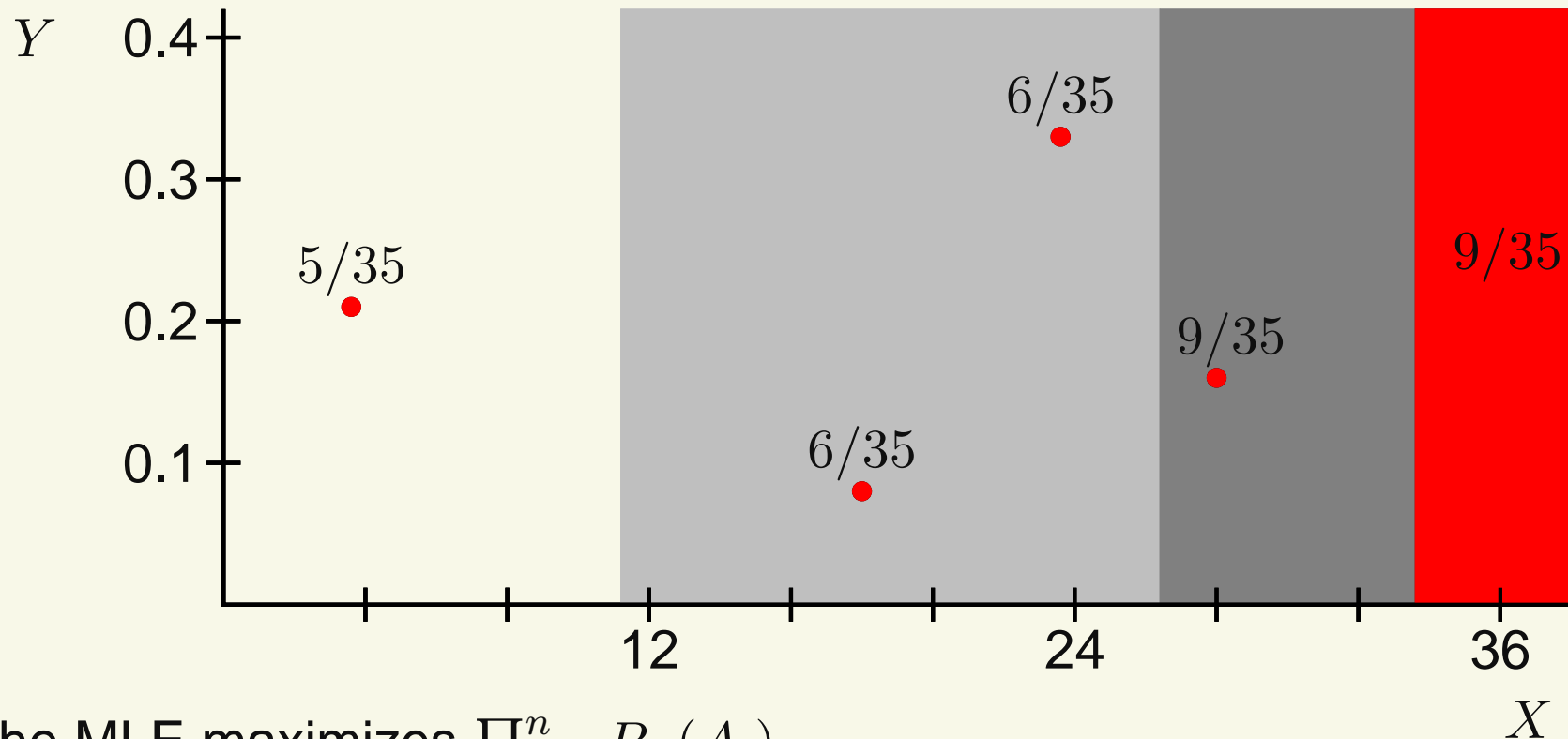
The MLE can only put mass in the maximal intersections.

The MLE may be non-unique.

$$\prod_{i=1}^n P_F(A_i) = p_1 p_2 p_3 p_4 (p_2 + p_3 + p_4 + p_5) (p_4 + p_5) p_5.$$

MLE for imputed data is the same as \hat{F}_n^ℓ .

The MLE can be viewed as a right endpoint imputation estimator



The MLE maximizes $\prod_{i=1}^n P_F(A_i)$.

The MLE can only put mass in the maximal intersections.

The MLE may be non-unique.

$$\prod_{i=1}^n P_F(A_i) = p_1 p_2 p_3 p_4 (p_2 + p_3 + p_4 + p_5) (p_4 + p_5) p_5.$$

MLE for imputed data is the same as \hat{F}_n^ℓ .

This indicates that \hat{F}_n^ℓ is biased.

MLE in terms of empirical processes

- Our likelihood is very similar to the likelihood for right censored data with continuous marks.

MLE in terms of empirical processes

- Our likelihood is very similar to the likelihood for right censored data with continuous marks.
- Hence, following Huang and Louis (1998), we can write:

- $1 - \widehat{F}_{X_n}^\ell(x) = \prod_{s \leq x} \{1 - \widehat{\Lambda}_{X_n}(ds)\}$

- $\widehat{F}_n^\ell(x, y) = \int_{s \leq x} \prod_{u < s} \{1 - \widehat{\Lambda}_{X_n}(du)\} \widehat{\Lambda}_n(ds, y)$

where

- $\widehat{\Lambda}_n(x, y) = \int_{[0, x]} \frac{\mathbb{V}_n(ds, y)}{1 - \mathbb{H}_n(s-)}$

- $\widehat{\Lambda}_{X_n}(x) = \widehat{\Lambda}_n(x, \infty) = \int_{[0, x]} \frac{\mathbb{V}_{X_n}(ds)}{1 - \mathbb{H}_n(s-)}$

and

- $\mathbb{H}_n(x) = \mathbb{P}_n 1\{U \leq x\}$

- $\mathbb{V}_n(x, y) = \mathbb{P}_n \Delta_+ 1\{U \leq x, Z \leq y\}$

- $\mathbb{V}_{X_n}(x) = \mathbb{V}_n(x, \infty) = \mathbb{P}_n \Delta_+ 1\{U \leq x\}$

and $U = \Delta_+ R + \Delta_{k+1} L$.

Limit of the MLE

- Uniform limits of \mathbb{H}_n , \mathbb{V}_n and \mathbb{V}_{X_n} (Glivenko-Cantelli):
 - $H(x) = P(U \leq x)$
 - $V(x, y) = P(\Delta_+1\{U \leq x, Z \leq y\})$
 - $V_X(x) = P(\Delta_+1\{U \leq x\})$
 - All three limits can be expressed in terms of F_0 and G .

Limit of the MLE

- Uniform limits of \mathbb{H}_n , \mathbb{V}_n and \mathbb{V}_{X_n} (Glivenko-Cantelli):
 - $H(x) = P(U \leq x)$
 - $V(x, y) = P(\Delta_+ 1\{U \leq x, Z \leq y\})$
 - $V_X(x) = P(\Delta_+ 1\{U \leq x\})$
 - All three limits can be expressed in terms of F_0 and G .
- Uniform limits of $\hat{\Lambda}_{X_n}$ and $\hat{\Lambda}_n$ (continuous mapping):
 - $\Lambda_{X_\infty} = \int_{[0,x]} \frac{V_X(ds)}{1-H(s-)}$
 - $\Lambda_\infty = \int_{[0,x]} \frac{V(ds,y)}{1-H(s-)}$

Limit of the MLE

- Uniform limits of \mathbb{H}_n , \mathbb{V}_n and \mathbb{V}_{X_n} (Glivenko-Cantelli):
 - $H(x) = P(U \leq x)$
 - $V(x, y) = P(\Delta_+ 1\{U \leq x, Z \leq y\})$
 - $V_X(x) = P(\Delta_+ 1\{U \leq x\})$
 - All three limits can be expressed in terms of F_0 and G .
- Uniform limits of $\widehat{\Lambda}_{X_n}$ and $\widehat{\Lambda}_n$ (continuous mapping):
 - $\Lambda_{X_\infty} = \int_{[0,x]} \frac{V_X(ds)}{1-H(s-)}$
 - $\Lambda_\infty = \int_{[0,x]} \frac{V(ds,y)}{1-H(s-)}$
- Uniform limits of $\widehat{F}_{X_n}^\ell$ and \widehat{F}_n^ℓ (continuous mapping):
 - $1 - F_{X_\infty}^\ell(x) = \prod_{s \leq x} \{1 - \Lambda_{X_\infty}(ds)\}$
 - $F_\infty^\ell(x, y) = \int_{s \leq x} \prod_{u < s} \{1 - \Lambda_{X_\infty}(du)\} \Lambda_\infty(ds, y)$

Limit of the MLE

- Uniform limits of \mathbb{H}_n , \mathbb{V}_n and \mathbb{V}_{X_n} (Glivenko-Cantelli):
 - $H(x) = P(U \leq x)$
 - $V(x, y) = P(\Delta_+ 1\{U \leq x, Z \leq y\})$
 - $V_X(x) = P(\Delta_+ 1\{U \leq x\})$
 - All three limits can be expressed in terms of F_0 and G .
- Uniform limits of $\widehat{\Lambda}_{X_n}$ and $\widehat{\Lambda}_n$ (continuous mapping):
 - $\Lambda_{X_\infty} = \int_{[0,x]} \frac{V_X(ds)}{1-H(s-)}$
 - $\Lambda_\infty = \int_{[0,x]} \frac{V(ds,y)}{1-H(s-)}$
- Uniform limits of $\widehat{F}_{X_n}^\ell$ and \widehat{F}_n^ℓ (continuous mapping):
 - $1 - F_{X_\infty}^\ell(x) = \prod_{s \leq x} \{1 - \Lambda_{X_\infty}(ds)\}$
 - $F_\infty^\ell(x, y) = \int_{s \leq x} \prod_{u < s} \{1 - \Lambda_{X_\infty}(du)\} \Lambda_\infty(ds, y)$
- Explicit formula for the limit of the MLE.

Necessary and sufficient conditions for consistency of $\widehat{F}_{X_n}^\ell$

- Marginal distribution of X :
 - Instantaneous hazard: $\lambda_{0X}(s) = \Lambda_{0X}(ds) = \frac{F_{0X}(ds)}{1 - F_{0X}(s-)}$
 - Cumulative hazard: $\Lambda_{0X}(x) = \int_{[0,x]} \frac{F_{0X}(ds)}{1 - F_{0X}(s-)}$
 - Then there is a one-to-one correspondence between F_{0X} and Λ_{0X} :
 $1 - F_{0X}(x) = \prod_{s \leq x} \{1 - \Lambda_{0X}(ds)\}$.

Necessary and sufficient conditions for consistency of $\widehat{F}_{X_n}^\ell$

- Marginal distribution of X :
 - Instantaneous hazard: $\lambda_{0X}(s) = \Lambda_{0X}(ds) = \frac{F_{0X}(ds)}{1 - F_{0X}(s-)}$
 - Cumulative hazard: $\Lambda_{0X}(x) = \int_{[0,x]} \frac{F_{0X}(ds)}{1 - F_{0X}(s-)}$
 - Then there is a one-to-one correspondence between F_{0X} and Λ_{0X} :
 $1 - F_{0X}(x) = \prod_{s \leq x} \{1 - \Lambda_{0X}(ds)\}$.
- Recall that $1 - F_{X_\infty}^\ell(x) = \prod_{s \leq x} \{1 - \Lambda_{X_\infty}(ds)\}$.
Hence, Λ_{X_∞} is the cumulative hazard of $F_{X_\infty}^\ell$.

Necessary and sufficient conditions for consistency of $\widehat{F}_{X_n}^\ell$

- Marginal distribution of X :
 - Instantaneous hazard: $\lambda_{0X}(s) = \Lambda_{0X}(ds) = \frac{F_{0X}(ds)}{1-F_{0X}(s-)}$
 - Cumulative hazard: $\Lambda_{0X}(x) = \int_{[0,x]} \frac{F_{0X}(ds)}{1-F_{0X}(s-)}$
 - Then there is a one-to-one correspondence between F_{0X} and Λ_{0X} :
 $1 - F_{0X}(x) = \prod_{s \leq x} \{1 - \Lambda_{0X}(ds)\}$.
- Recall that $1 - F_{X_\infty}^\ell(x) = \prod_{s \leq x} \{1 - \Lambda_{X_\infty}(ds)\}$.
Hence, Λ_{X_∞} is the cumulative hazard of $F_{X_\infty}^\ell$.
- The MLE $\widehat{F}_{X_n}^\ell$ is consistent for F_{0X} on $[0, \tau]$ if and only if for all $x \in [0, \tau]$:

$$\Lambda_{X_\infty}(x) \equiv \int_{[0,x]} \frac{V_X(ds)}{1 - H(s-)} = \int_{[0,x]} \frac{F_{0X}(ds)}{1 - F_{0X}(s-)} \equiv \Lambda_{0X}(x).$$

Necessary and sufficient conditions for consistency of \widehat{F}_n^ℓ

- Bivariate distribution of (X, Y) (Huang and Louis, 1998):
 - Mark-specific instantaneous hazard: $\Lambda_0(ds, y) = \frac{F_0(ds, y)}{1 - F_{0X}(s-)}$
 - Mark-specific cumulative hazard: $\Lambda_0(x, y) = \int_{[0, x]} \frac{F_0(ds, y)}{1 - F_{0X}(s-)}$
 - Then there is a one-to-one correspondence between F_0 and Λ_0 :
$$F_0(x, y) = \int_{s \leq x} \prod_{u < s} \{1 - \Lambda_{0X}(du)\} \Lambda_0(ds, y).$$

Necessary and sufficient conditions for consistency of \widehat{F}_n^ℓ

- Bivariate distribution of (X, Y) (Huang and Louis, 1998):
 - Mark-specific instantaneous hazard: $\Lambda_0(ds, y) = \frac{F_0(ds, y)}{1 - F_{0X}(s-)}$
 - Mark-specific cumulative hazard: $\Lambda_0(x, y) = \int_{[0, x]} \frac{F_0(ds, y)}{1 - F_{0X}(s-)}$
 - Then there is a one-to-one correspondence between F_0 and Λ_0 :
$$F_0(x, y) = \int_{s \leq x} \prod_{u < s} \{1 - \Lambda_{0X}(du)\} \Lambda_0(ds, y).$$
- Recall that $F_\infty^\ell(x, y) = \int_{s \leq x} \prod_{u < s} \{1 - \Lambda_{X\infty}(du)\} \Lambda_\infty(ds, y)$.
Hence, Λ_∞ is the mark-specific cumulative hazard of F_∞^ℓ .

Necessary and sufficient conditions for consistency of \widehat{F}_n^ℓ

- Bivariate distribution of (X, Y) (Huang and Louis, 1998):
 - Mark-specific instantaneous hazard: $\Lambda_0(ds, y) = \frac{F_0(ds, y)}{1 - F_{0X}(s-)}$
 - Mark-specific cumulative hazard: $\Lambda_0(x, y) = \int_{[0, x]} \frac{F_0(ds, y)}{1 - F_{0X}(s-)}$
 - Then there is a one-to-one correspondence between F_0 and Λ_0 :
$$F_0(x, y) = \int_{s \leq x} \prod_{u < s} \{1 - \Lambda_{0X}(du)\} \Lambda_0(ds, y).$$
- Recall that $F_\infty^\ell(x, y) = \int_{s \leq x} \prod_{u < s} \{1 - \Lambda_{X\infty}(du)\} \Lambda_\infty(ds, y)$.
Hence, Λ_∞ is the mark-specific cumulative hazard of F_∞^ℓ .
- The MLE \widehat{F}_n^ℓ is consistent for F_0 on $[0, \tau] \times \mathbb{R}$ if and only if for all $x \in [0, \tau]$ and $y \in \mathbb{R}$:

$$\Lambda_\infty(x, y) \equiv \int_{[0, x]} \frac{V(ds, y)}{1 - H(s-)} = \int_{[0, x]} \frac{F_0(ds, y)}{1 - F_{0X}(s-)} \equiv \Lambda_0(x, y).$$

Corollaries

- MLE is consistent if and only if a certain relation between F_0 and G holds. This will typically not be the case.

Corollaries

- MLE is consistent if and only if a certain relation between F_0 and G holds. This will typically not be the case.
- Corollary 1: Let X be subject to current status censoring ($k = 1$). Then the MLE \widehat{F}_{Xn}^ℓ is inconsistent for **any** choice of continuous distributions F_{0X} and G .

Corollaries

- MLE is consistent if and only if a certain relation between F_0 and G holds. This will typically not be the case.
- Corollary 1: Let X be subject to current status censoring ($k = 1$). Then the MLE $\widehat{F}_{X_n}^\ell$ is inconsistent for **any** choice of continuous distributions F_{0X} and G .
- Corollary 2: Let X be subject to interval censoring case k , and let the elements T_1, \dots, T_k of \mathbf{T} be the order statistics of k independent uniform random variables on $[0, \theta]$. Then MLE becomes consistent as $k \rightarrow \infty$.

Corollaries

- MLE is consistent if and only if a certain relation between F_0 and G holds. This will typically not be the case.
- Corollary 1: Let X be subject to current status censoring ($k = 1$). Then the MLE $\widehat{F}_{X_n}^\ell$ is inconsistent for **any** choice of continuous distributions F_{0X} and G .
- Corollary 2: Let X be subject to interval censoring case k , and let the elements T_1, \dots, T_k of \mathbf{T} be the order statistics of k independent uniform random variables on $[0, \theta]$. Then MLE becomes consistent as $k \rightarrow \infty$.
- Corollary 2 can also be understood by viewing the MLE as a right endpoint imputation estimator.

Repairing inconsistency by discretizing the marks

- Discretize marks:
 - Let $K > 0$ and define grid $-\infty \equiv y_0 < y_1 < \cdots < y_K < y_{K+1} \equiv \infty$
 - Define competing risk $C \in \{1, \dots, K + 1\}$:

$$y_{j-1} < Y \leq y_j \quad \Rightarrow \quad C = j, \quad j = 1, \dots, K + 1.$$

- We can observe C for all observations with an observed mark.
- So we can transform the observations $W = (\mathbf{T}, \mathbf{\Delta}, Z)$ into $W^* = (\mathbf{T}, \mathbf{\Delta}, Z^*)$, where $Z^* = \Delta_+ C$.

Repairing inconsistency by discretizing the marks

- Repaired MLE:
 - Consider MLE \tilde{F}_n for distribution of (X, C) based on i.i.d. observations W_1^*, \dots, W_n^* .
 - \tilde{F}_n is MLE for interval censored data with competing risks.

Repairing inconsistency by discretizing the marks

- Repaired MLE:
 - Consider MLE \tilde{F}_n for distribution of (X, C) based on i.i.d. observations W_1^*, \dots, W_n^* .
 - \tilde{F}_n is MLE for interval censored data with competing risks.
- Consistency of the repaired MLE:
 - Consistency of \tilde{F}_n follows from empirical process theory (see, e.g., Maathuis (2006)).
 - We can consistently estimate $P(X \leq x, C \leq j) = P(X \leq x, Y \leq y_j) = F_0(x, y_j)$ for $j = 1, \dots, K$.
 - We can consistently estimate F_0 for y on the grid.

Simulation to illustrate behavior of MLE and repaired MLE

- Example 1:

- $X \sim \text{Unif}(0, 1)$, $Y \sim \text{Exp}(1)$, X and Y independent
- $T \sim \text{Unif}(0, 0.5)$

- Example 2:

- $X \sim \text{Unif}(0, 1)$, $Y|X \sim \text{Exp}(X + 0.5)$
- $T \sim \text{Unif}(0, 1)$

- Example 3:

- $X \sim \text{Unif}(0, 2)$, $Y = X$
- $T_1 \sim \text{Unif}(0, 1)$, $T_2 \sim \text{Unif}(0, 2)$

- Example 4:

- (X, Y) uniform over $\{(x, y) : 0 \leq x \leq y \leq 1\}$
- (T_1, T_2) discrete with $G(0.25, 0.5) = 0.3$, $G(0.25, 0.75) = 0.3$ and $G(0.5, 0.75) = 0.4$

Simulation to illustrate behavior of MLE and repaired MLE

- Example 1:

- $X \sim \text{Unif}(0, 1)$, $Y \sim \text{Exp}(1)$, X and Y independent
- $T \sim \text{Unif}(0, 0.5)$

- Example 2:

- $X \sim \text{Unif}(0, 1)$, $Y|X \sim \text{Exp}(X + 0.5)$
- $T \sim \text{Unif}(0, 1)$

- Example 3:

- $X \sim \text{Unif}(0, 2)$, $Y = X$
- $T_1 \sim \text{Unif}(0, 1)$, $T_2 \sim \text{Unif}(0, 2)$

- Example 4:

- (X, Y) uniform over $\{(x, y) : 0 \leq x \leq y \leq 1\}$
- (T_1, T_2) discrete with $G(0.25, 0.5) = 0.3$, $G(0.25, 0.75) = 0.3$ and $G(0.5, 0.75) = 0.4$

- One simulation with $n = 10,000$. Use $K = 20$ for \tilde{F}_n .

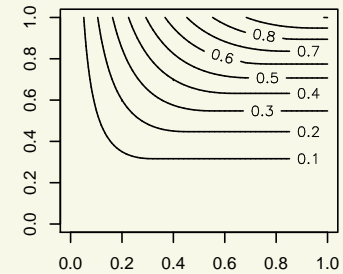
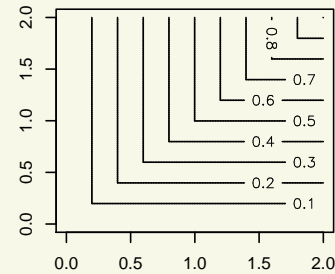
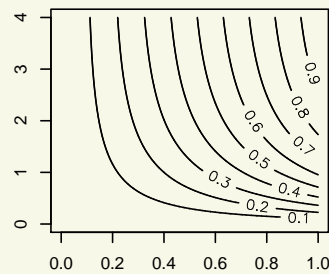
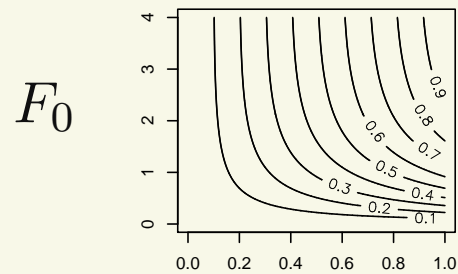
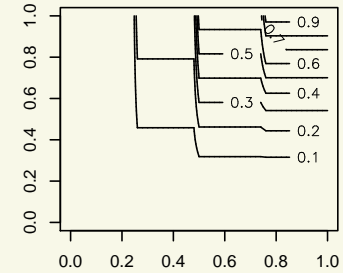
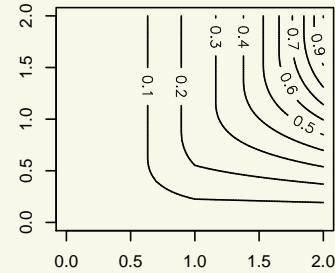
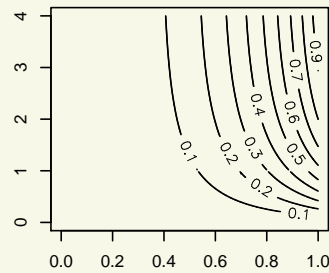
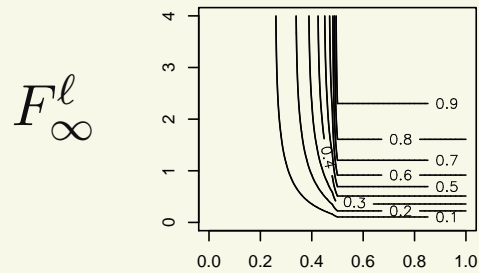
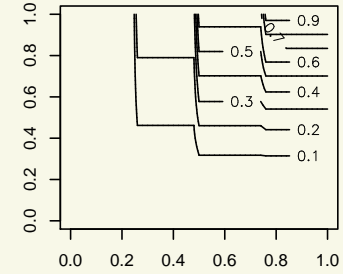
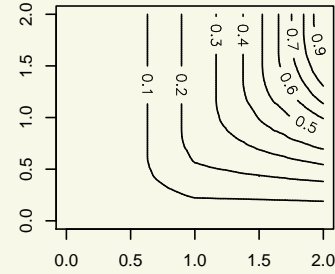
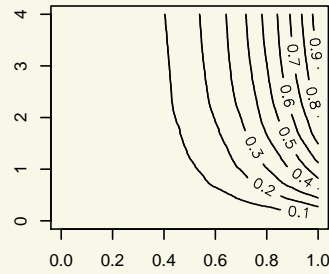
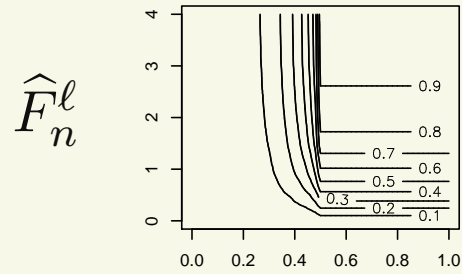
Estimating joint distribution $F_0(x, y)$

Ex. 1

Ex. 2

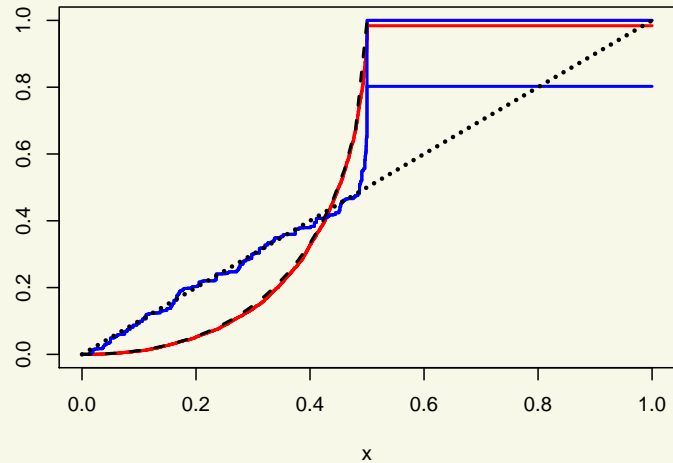
Ex. 3

Ex. 4

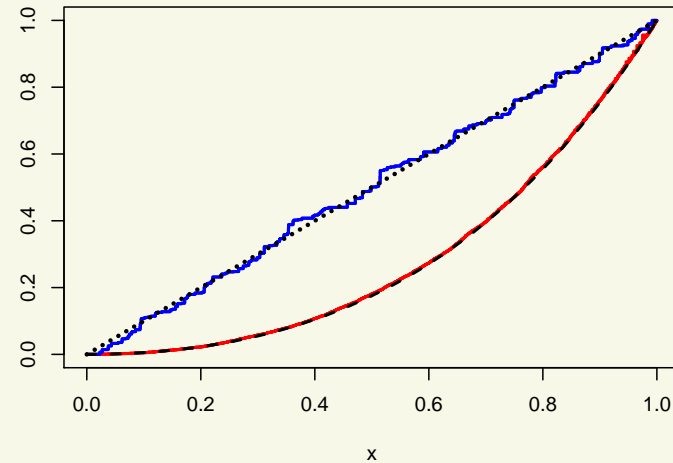


Estimating marginal distribution of X

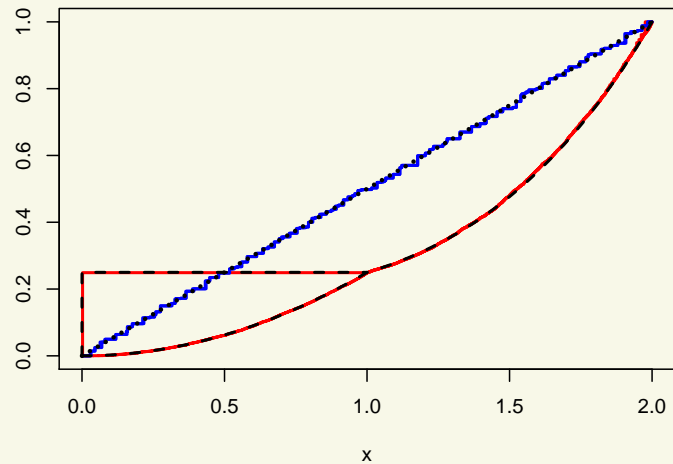
Example 1



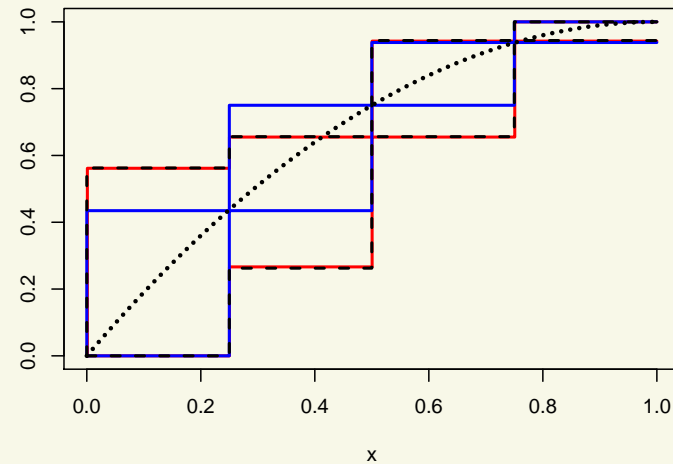
Example 2



Example 3



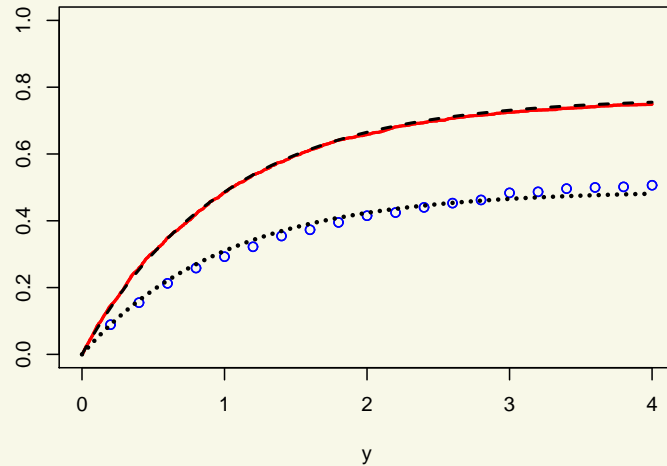
Example 4



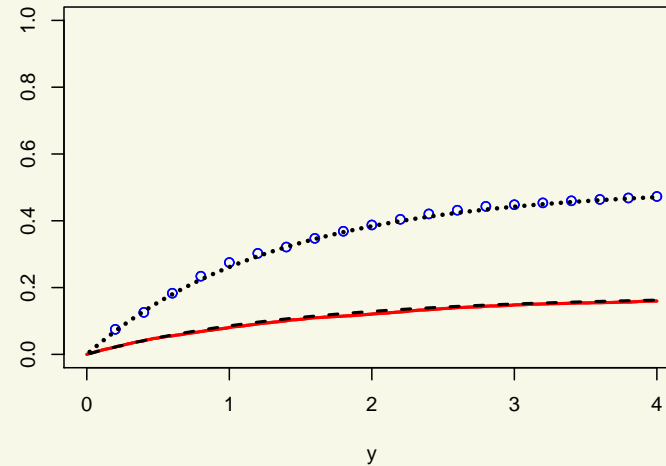
Legend: dotted= $F_{0X}(\cdot)$, red= $\hat{F}_{X_n}(\cdot)$, dashed= $F_{X_\infty}(\cdot)$, blue= $\tilde{F}_{X_n}(\cdot)$

Estimating $F_0(x_0, y)$ for fixed x_0

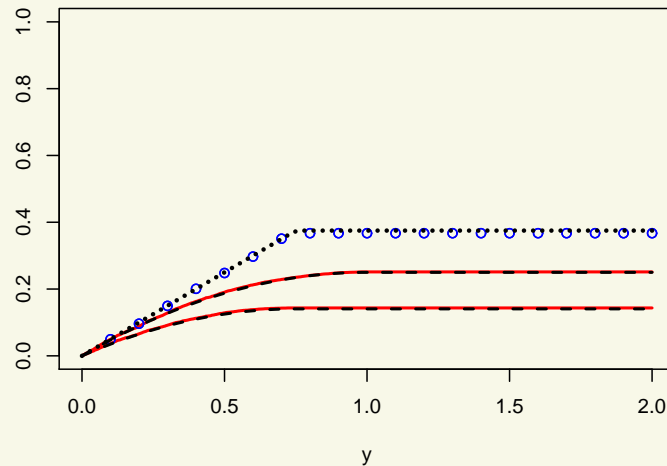
F(0.49, y), Example 1



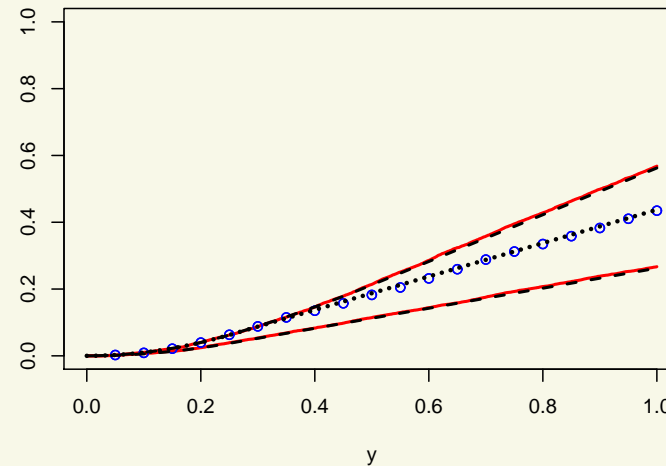
F(0.5, y), Example 2



F(0.75, y), Example 3



F(0.25, y), Example 4



Legend: dotted= $F_0(x_0, \cdot)$, red= $\hat{F}_n(x_0, \cdot)$, dashed= $F_\infty(x_0, \cdot)$, blue= $\tilde{F}_n(x_0, \cdot)$

Summary

- We studied the MLE for the distribution of an interval censored survival time and a continuous mark.
- We derived an explicit formula for the limit of the MLE.
- We showed that the MLE is inconsistent in general.
- Inconsistency can be repaired by discretizing the marks.
- Simulation results indicate that:
 - The MLE can be off by a lot
 - The repaired MLE performs well

Discussion

- Relation to previous work:
 - Our results confirm that line segments are a warning sign for consistency problems.
 - In Maathuis (2003), inconsistency of the MLE was caused by non-uniqueness of the MLE. That is not the case in the current model. The MLE is typically inconsistent even if the limit is unique.

Discussion

- Relation to previous work:
 - Our results confirm that line segments are a warning sign for consistency problems.
 - In Maathuis (2003), inconsistency of the MLE was caused by non-uniqueness of the MLE. That is not the case in the current model. The MLE is typically inconsistent even if the limit is unique.
- Extension to more complicated models:
 - Extension to mixed case interval censoring is straightforward.
 - If marks are missing for some observations with $\Delta_+ = 1$, then there is no closed form available. Hence, in this case consistency is still open problem, but simulation results point to inconsistency.

References

- M.G. HUDGENS, M.H. MAATHUIS AND P.B. GILBERT (2007). Nonparametric estimation of the joint distribution of a survival time subject to interval censoring and a continuous mark variable. *Biometrics* **63** 372–380.
- M.H. MAATHUIS AND J.A. WELLNER (2007). Inconsistency of the MLE for the joint distribution of interval censored survival times and continuous marks. *Scandinavian Journal of Statistics*, *accepted*.

See <http://www.stat.washington.edu/marloes>

Thanks!