

第2章： 各統計手法の批判的検討

1. 帰無仮説検定法
2. 情報理論法
3. ベイズ法

まず例題で比較
あとで個別に検討

例題：コアラ袋内児の性比

12頭の母親を検査：
3頭♂、9頭♀
(コアラは一産一子)

何が言えるか？



帰無仮説検定

帰無仮説: 性比 $r=0.5$

対立仮説: 性比 $r < 0.5$

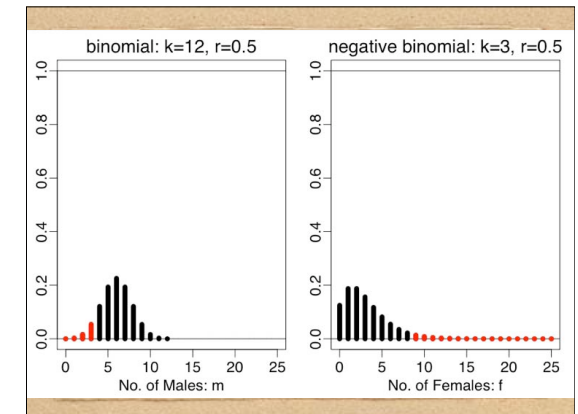
性比: 全体に占める♂の割合

データ抽出方法： 「打ち切り基準 stopping rule」

- 1) 母親の数が12頭になるまで抽出
結果として♂子3頭
- 2) ♂子の数が3頭になるまで抽出
結果として♀子9頭

Box 2.1 各打ち切り基準に対する 帰無仮説検定

- 1) 母親数12で抽出: ♂子数3
比率 $r = 0.5$ の二項分布: $x \leq 3$
 $p = 0.073$: 帰無仮説は棄却されず
- 2) ♂子数3で抽出: ♀子数9
比率 $r = 0.5$ の負の二項分布: $x \geq 9$
 $p = 0.033$: 帰無仮説を棄却
対立仮説 ($r < 0.5$) を受容



帰無仮説検定のまとめ

結論がデータ抽出法に依存 ???

原因
観察データだけでなく、それより末端側
にある統計値に依存 ???

帰無仮説検定のまとめ

結論がデータ抽出法に依存

異議あり！

結果同じでも、抽出法が違えば違うデータ
結論が違ってもおかしくない

原因
観察データだけでなく、それより末端側
にある統計値に依存 ???

異議あり！

末端から確率累積は当該データが確率分布の中で
どこにあるかを表す手段に過ぎない

情報理論法

パラメータ決定に最尤推定を使用

観察データを得る確率(尤度)を最大化するもの

最大尤度を含む信頼区間を推定

各打ち切り基準に対する 最尤推定

尤度: 性比 r における条件付き確率

1) 母親数12で抽出: ♂子3

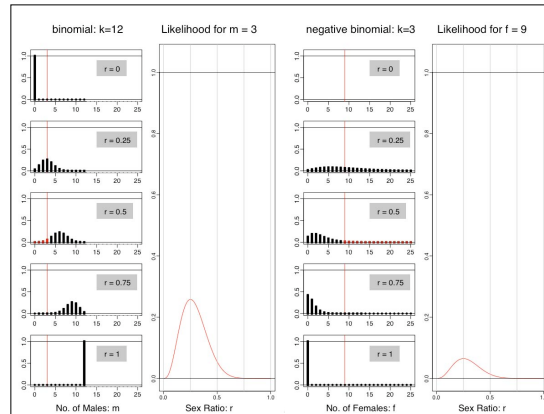
$$\Pr(x|r) = {}_{12}C_x r^x (1-r)^{(12-x)} = {}_{12}C_3 r^3 (1-r)^9$$

2) ♂子数3で抽出: ♀子9

$$\Pr(x|r) = {}_{x+(3-1)}C_x r^3 (1-r)^x = {}_{11}C_9 r^3 (1-r)^9$$

どちらも 尤度 $L = r^3(1-r)^9$

$r = 0.25$ で 最尤値 $L_{\max} = 0.001173198$



信頼区間

$r = 0.25$

最尤値 $L_{\max} = 0.001173$

自由度1のカイニ乗分布の95%点: 3.84

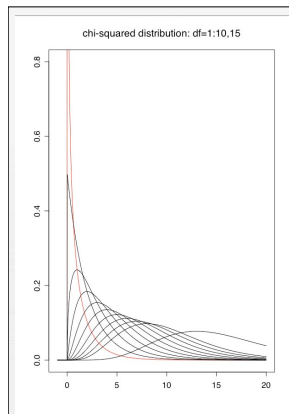
$\exp(-3.84/2) = 0.1465$

$0.1465 L_{\max} = 0.000172$

$L = r^3(1-r)^9 = 0.000172$ を解く

$r = 0.069$ (下限), 0.528 (上限)

一体何をやっているか続くスライド2枚で解説

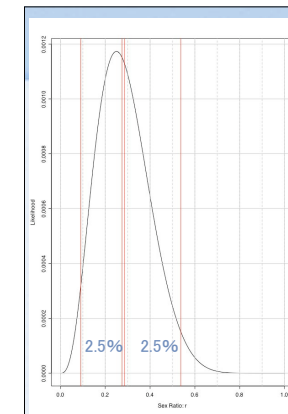
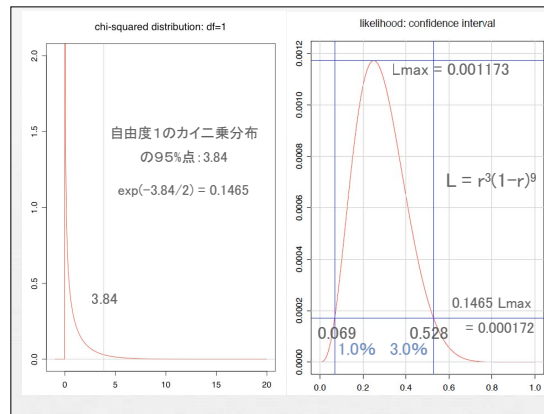


カイニ乗分布とは

標準正規分布をとる
変数k個の二乗和が
とる確率分布

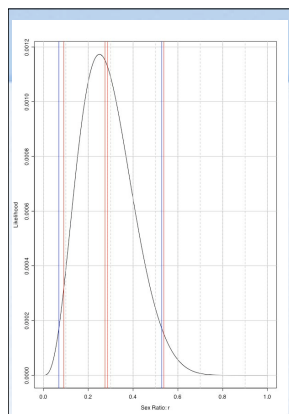
$$\sum_{i=1}^k x_i^2$$

kは自由度



ベイズ法

最頻値 $r = 0.250$
中央値 $r = 0.275$
平均値 $r = 0.286$
95%**確信**区間:
 $r = 0.091-0.537$
credible interval



情報理論法

ベイズ法

95%**確信**区間:

$r = 0.091-0.537$

credible interval

情報理論法

95%**信頼**区間

$r = 0.069-0.528$

confidence interval

比率の点推定、区間推定

・標本比率: $r = 0.25$
母集団比率 π の点推定値

・ r の標準誤差: $SE = \sqrt{r(1-r)/N} = 0.125$

母比率 π の95%信頼区間:

$r - 1.96 SE < \pi < r + 1.96 SE$

$0.005 < \pi < 0.0495$

頻度主義法

ベイズ法

95%**確信**区間:

$r = 0.091-0.537$

尤度法

95%**信頼**区間

$r = 0.069-0.528$

頻度主義

95%**信頼**区間

$r = 0.005-0.495$

ベイズ法のまとめ

仮説検定法との比較
結論が違う ???

情報理論法との比較

事前分布に情報なし: 似た推定 ?
事前分布に情報あり: ベイズ法に利

ベイズ法のまとめ

仮説検定法との比較
結論が違う ???

異議あり: 問い方が違う; 結論が違うのは当然
仮説検定法: 「この仮説で良いか」
情報理論法・ベイズ法: 「どの仮説がよさそうか」

情報理論法との比較

事前分布に情報なし: 似た推定 ?
事前分布に情報あり: ベイズ法に利
疑問: なぜ「同じ」でない?
なぜ区間推定異なる?

例題終了

休憩

次、個別に検討

帰無仮説検定の手順

1. 帰無・対立仮説の設定
2. データ収集
3. 帰無仮説下
観察データとそれより末端側のデータを得る確率 p を計算
4. 確率 $p < \text{基準値}$ (通常 $p=0.05$)
帰無仮説下で観察データはありそうもない
帰無仮説はありそうもない、対立仮説を採用
5. 確率 $p > \text{基準値}$: 帰無仮説、棄却できず。
帰無仮説を採用

帰無仮説検定:
帰無仮説を不適切とは言えるが、正しいとは言えない ???

ばかげた仮説 silly null

例: Mullan Crain et al. 2004.

帰無仮説:

塩、汽、淡水の沼地は種多様度が同じ

対立仮説:

塩分濃度が違えば種多様度が異なる

帰無仮説: 端から正しいそうもない
棄却が深い理解へつながらない

なぜ silly nulls?

非ゼロ効果の帰無仮説設定が困難
沼地の例:

具体的な違いに関する仮説なし

塩分濃度による種多様度予測理論:
あれば有効な帰無仮説に
棄却の意味: 理論が無効

生態学: 厳密な理論のぞめない

ポッパの反証主義

「理論は仮説に過ぎず、仮説は証明不可能。
だが棄却は可能。」
(Karl Popper: ウィーンの科学哲学者)

ばかげた帰無仮説の棄却:
ポッパ主義の基準を満たさず
ポッパ主義:
帰無仮説検定なしでも達成可能

帰無仮説検定のp値

帰無仮説下で、
観察データとそれより末端側のデータを得る確率 ???

観察されないデータが推定に影響: 合理的な懸念 ???

帰無仮説検定のp値

帰無仮説下で、
観察データとそれより末端側のデータを得る確率 ???

異議あり: p 値は当該データが確率分布の中で
どこに位置するかを示すだけ

観察されないデータが推定に影響: 合理的な懸念 ???

異議あり: 観察されないデータが影響するのではなく、
当該データが確率分布中のどこにあるかが影響するだけ

p値小による帰無仮説棄却

1. p値: 帰無仮説の正しさを示すと誤解
2. 有意水準: 根拠なし
通常 $p < 0.05$: 危険率 α (有意水準)
タイプIエラー: 正しい帰無仮説を棄却する確率
3. 検出力: 考慮せず
 $\text{Power (検出力)} = 1 - \text{タイプIIエラー}$
タイプIIエラー: 不当な帰無仮説を受容する確率
 Power : 不当な仮説を(正しく)棄却する確率

帰無仮説の棄却

検討なしに対立仮説を受容
対立仮説: 帰無仮説より不適切な可能性も

Box 2.3 Sally Clark事件

幼子2人相次ぎ死亡、原因不明。殺人で有罪に
帰無仮説: 殺していない、**対立仮説**: 殺した

帰無仮説下で相次ぐ死亡の確率: 7300万分の1
対立仮説(殺した)を受容

帰無仮説の棄却(p値小):
対立仮説の正しさを意味せず

p値大による帰無仮説受容

p値大: 帰無仮説真を示すと誤解
有意差なし: 帰無仮説真と判断
検出力なしで判断不可能

検出力計算:

検出したい差とデータのばらつきを指定必要
小さい差: 検出しにくい
ばらつきを過小推定: 検出力下がる

帰無仮説検定法のまとめ

- 問題: 主に利用法にあり
1. 「ばかげた帰無仮説」の使用
 2. p値の意味を誤解
 3. 検出力を無視

根本問題

- A. 観察されないデータが結果に影響 ???
- B. 対立仮説の妥当性検討されず

帰無仮説検定法のまとめ

- 問題: 主に利用法にあり
1. 「ばかげた帰無仮説」の使用
 2. p値の意味を誤解
 3. 検出力を無視

根本問題

- A. 観察されないデータが結果に影響 ???
- B. 対立仮説の妥当性検討されず

**根本問題: 問い方につける
仮説二つ、片方だけ検定**

情報理論法の手順

1. 仮説として複数モデルを選択
2. データ収集
3. 各モデルを評価
各モデルの情報損失量を推定
損失量最小のモデルを選択
4. 各モデルを情報損失の少なさで重み付け
要因の効果を平均で予測

複数のモデル

帰無仮説検定: 二つの仮説(帰無と対立)

情報理論法: 複数の仮説を同時に検討

モデルとは仮説の具体化

観察データとパラメータを関連づける関数

分析目的:

「簡素で最も効率よく」現象を説明するモデル

most parsimonious

例: ケワタガモ生存率

仮説

1. 鉛量のみ
2. 鉛量と場所が相加的
3. 鉛量と場所が交互作用
4. 場所のみ
5. 無関係

Grand et al. 1998



Spectacled Eider, メガネケワタガモ

1, 2, 4, 5は「ばかげた仮説」で、3がベストかい?

No! 簡素主義に反

帰無仮説検定法との違い

帰無仮説検定:

仮説棄却時、別の仮説を検討

決定的な違い

情報理論法:

情報損失量を最小化

帰無仮説検定:

タイプIエラーに依存

基準に根拠のない0.05を使用

情報損失量を計算

AIC: Akaikeの情報量基準

$AIC = -2\ln(L) + 2k$: L (尤度)、 k (パラメータ数)

モデル適合度とパラメータ数から計算
簡素主義を内包

Akaikeとは赤池弘次(元文部省統計数理研究所長)

適合度低で複雑: AIC大

最良モデル: 情報損失量(AIC)最小

ケワタガモの例:

鉛のみ仮説が最良、鉛・場所相加仮説が次

AICの応用

AIC: 大小関係が重要

最小値との差分で表現

$\Delta i = AIC_i - AIC_{min}$

AICの差分:

Akaikeの重み w_i に変換

$w_i = \exp(-\Delta i/2) / \sum (r=1 \sim R) \exp(-\Delta r/2)$

ケワタガモの例

鉛を1要因とするモデルの w_i の総和 (0.998)

鉛が死亡率に影響するとする強い根拠

モデル平均

要因の効果推定

最良モデルのみ検討は最良にあらず

各モデルの平均をとる

ケワタガモの例:

鉛: 生存率を0.335低下、標準誤差0.125

情報理論法:

どのモデルが最良か、確定不可能と認識

一つのモデルに終始する危険を回避

情報理論法の利点

帰無仮説検定法と比べて

1. 観察値より末端側データに影響されない ???
2. 各仮説の妥当性を相対的に同時に評価
3. 要因の効果を各仮説の相対的重みで推定

情報理論法: 利用上の問題

1. 「ばかげた仮説」
2. Akaikeの重み w_i を仮説真の確率と想定
3. 最良モデルのみ採用、他を無視
4. 調査計画の妥当性を検討せず
5. 候補モデル絞り込み、任意のAIC差分(Δ)基準値を使用
6. 最良モデル決定で、効果の大きさを無視

ベイズ法の手順

1. 仮説として複数モデルを使用
2. 各仮説に事前確率を適用
3. データ収集
4. 事前確率と観察データを併せて事後予測

複数仮説の使用

仮説数に制約なし:

ばかげた仮説使用に制約なし:

ベイズ法:

各仮説の確率を事前に付与

ばかげた仮説は確率小

事前確率

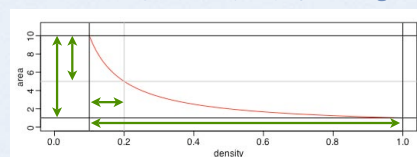
等確率付与の問題

・離散パラメータ: 仮説の数に確率が依存

・連続パラメータ: 使用単位により意味が変化

例: なわばりの密度と面積

両単位の関係(逆数): 非線形 Fig. 2.2



事前情報から事前確率へ

事前確率付与における主観性:

人、問題、動機、主義に依存

科学における主観性: 不可避、でも最小に

注意深く、厳密に、説得力を持って!

どうして良いか分からない場合

主観の明示的な導入を!

科学の役割:

(主観的) 想定を根拠に基づき論理的に更新

事前確率分布の使用

ベイズ法の難点であり、売りでも。

研究論文: 先行研究の結果を必ず参照

ベイズ法:

既存の知見を参照するための形式的枠組み

ベイズの法則を利用

より厳密・客観的に既存の知見を使用

ベイズの法則

観察データを得た後の各仮説の確率

当仮説の事前確率 と

当仮説下でそのデータを得る確率 の積に比例

各仮説の事前想定を論理的に更新

問題: どう事前確率を与えるか

困難。だが、

事前情報があれば、使わないのは不自然

パラメータの確率分布

パラメータ: ランダム変数か?

必ずしもそうでない

未知の固定値で確率的にのみ表現可能

ベイズ法の利点

1. 適応条件がゆるい
分散分析: 等分散が条件
ベイズ推定: 不等分散可
回帰分析でも、等分散性を前提とせず
 2. 階層的効果の取り込みが比較的簡単
偽反復問題に対処容易
- 研究デザインを分析に合わせるのではなく、
研究デザインに合わせた分析が可
- 出所の異なる多くのデータに対応可、
多様な統計モデルを使用可

問題点

計算負荷が高い(MCMCなど)
ソフトに必要な関数があることが必要

時間が解決

効果の大きさの推定

あるべき研究上の設問:

沼地の種多様度: 塩分濃度で

×「変わるのか?」

◎「どう変わるか?」

効果の大きさの推定とその精度の尺度

ベイズ法に限らず、どんな統計分析でも要

精度の尺度: 区間推定

頻度主義の信頼区間

ベイズ推定の確信区間

区間推定の意味

生態学的に意味ある解釈可能

検出力を区間幅で表現可能

単なる有意差の有無は無意味

効果の大きさ、推定値の精度、変量間の関係

生態学的に意味ある結果

Sally Clark事件

幼子2人相次ぎ死亡、原因不明。殺人で有罪に

仮説検定 帰無仮説: 彼女は殺していない
相次ぐ死亡の確率: 7300万分の1
対立仮説: 彼女が殺した

対立仮説を支持

情報理論 仮説1 : 彼女は殺していない
相次ぐ死亡の尤度: 30万分の1 (情報量) 仮説2 : 彼女が殺した
相次ぐ死亡の尤度: 1.0

ベイズ法 観による子殺しの確率を考慮
仮説1 : 彼女は殺していない 事前確率: $1 - 3.7 \times 10^{-7}$
相次ぐ死亡の事後確率: 0.1
仮説2 : 彼女が殺した 事前確率: 3.7×10^{-7}
相次ぐ死亡の事後確率: 0.9

第2章のまとめ

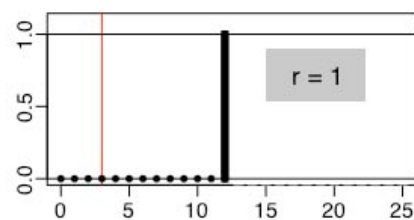
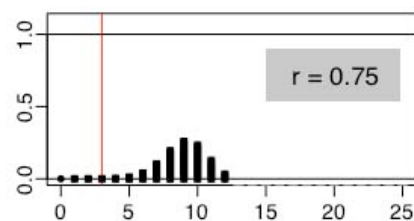
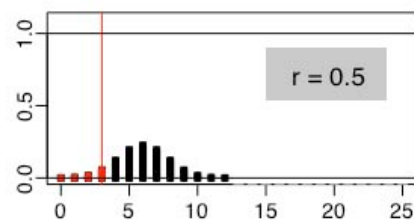
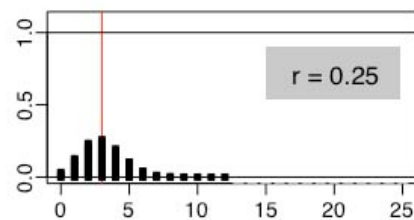
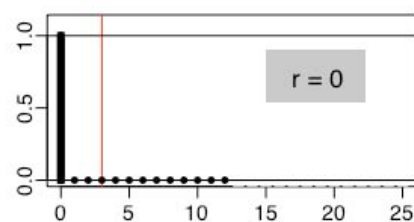
	問い方	事前情報
仮説検定法	これでよいか	×
情報理論法	どれが	×
ベイズ法	よさそうか	○

binomial: $k=12$

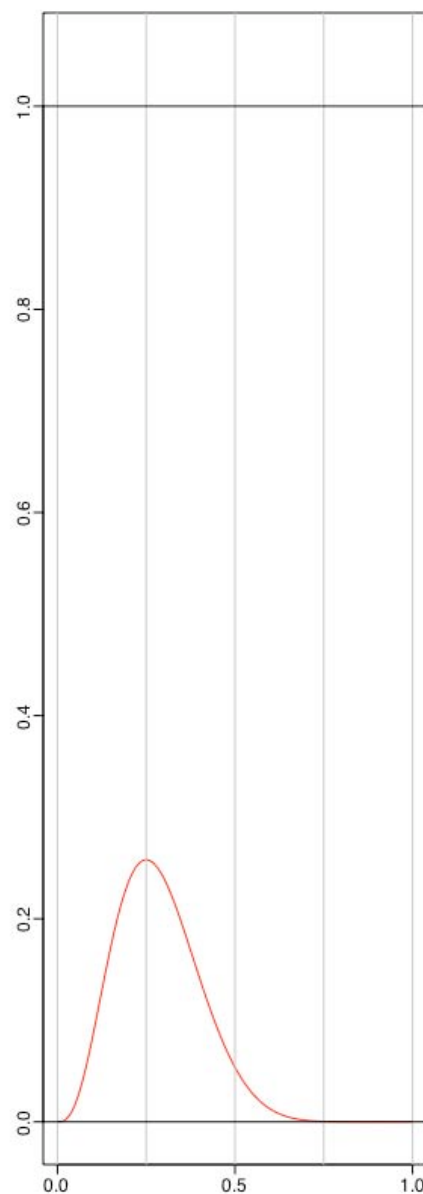
Likelihood for $m = 3$

negative binomial: $k=3$

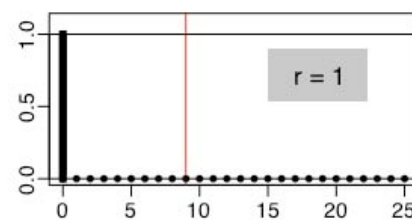
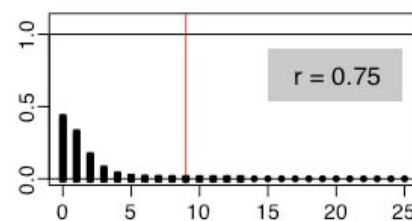
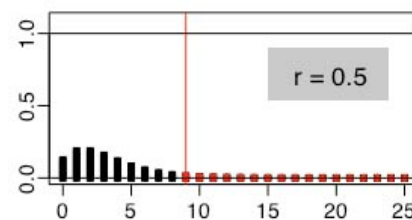
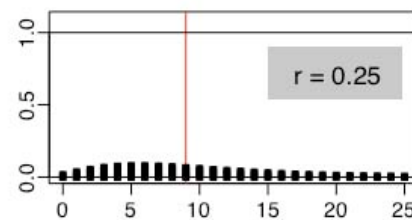
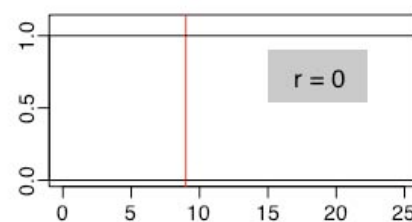
Likelihood for $f = 9$



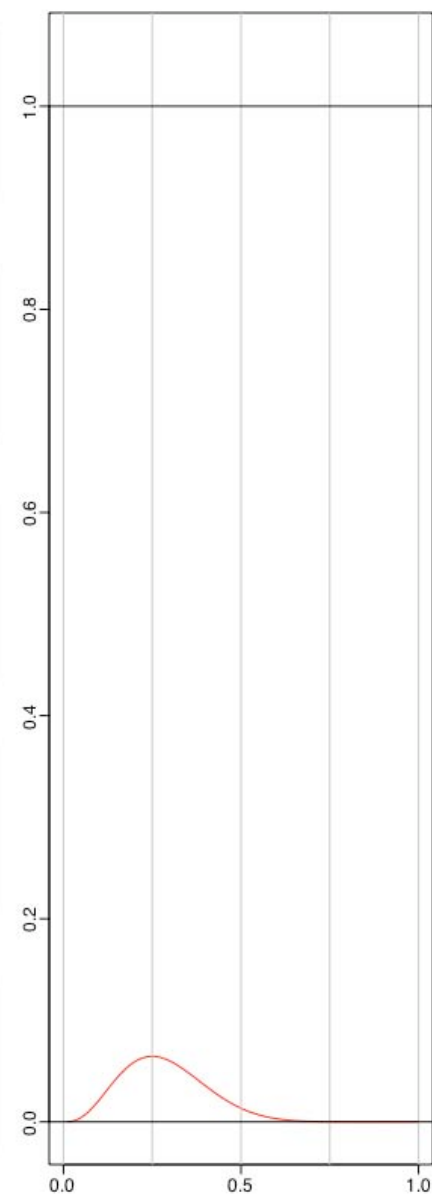
No. of Males: m



Sex Ratio: r



No. of Females: f



Sex Ratio: r