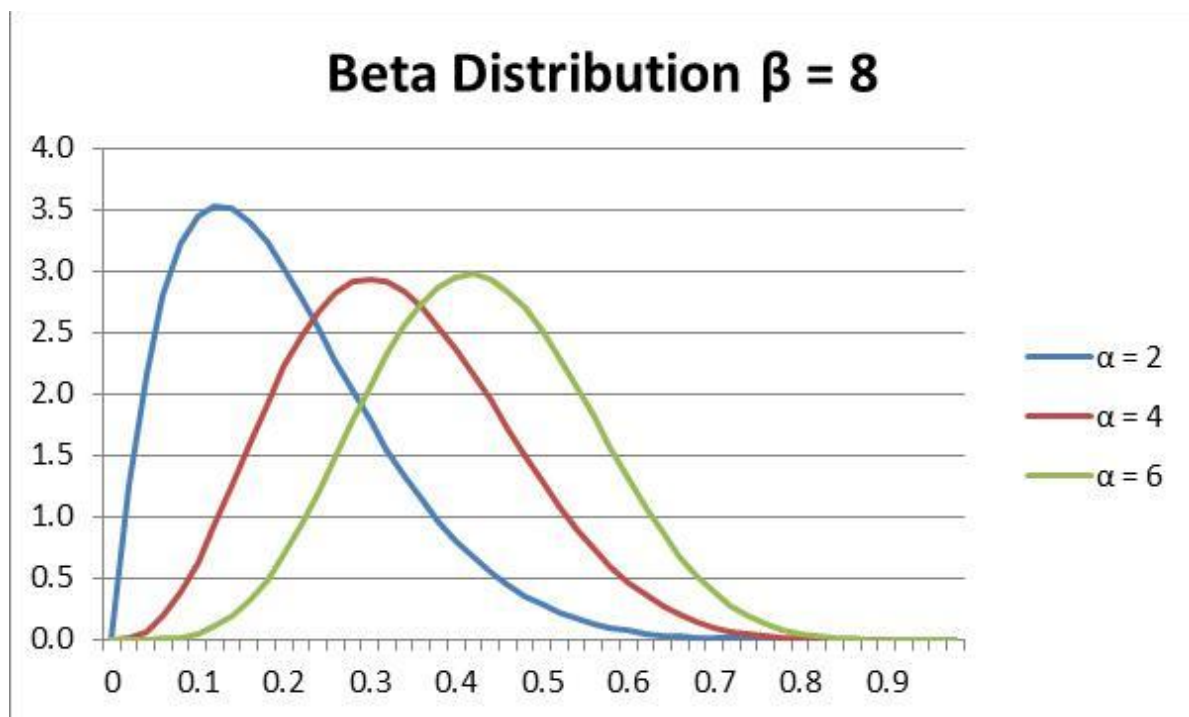


[Stat] Memahami Distribusi Beta (Beta Distribution) Dengan Contoh Nyata (Statistik Permainan Mobile Legend)

Suberlin Sinaga



Pada dasarnya banyak pelajar statistika pemula mengetahui beberapa jenis distribusi seperti normal, binomial, dan atau seragam.

Banyak pelajar pemula dalam ilmu statistik memiliki pemahaman yang baik akan distribusi normal, binomial, dan atau seragam. Namun sayangnya, banyak dari pemula yang tidak memiliki pemahaman yang memadai terhadap distribusi beta.

Hal ini disebabkan karena distribusi seperti distribusi normal, binomial, dan atau distribusi seragam dijelaskan dengan sangat baik, lengkap dan jelas dalam berbagai literatur. Sementara distribusi beta hanya akan dibahas pada kondisi-kondisi tertentu dan dengan penjelasan yang sulit dimengerti.

Berkaca pada fenomena ini, disini saya berusaha untuk mencoba menjelaskan apa itu distribusi beta dan bagaimana penggunaannya dalam kehidupan nyata.

Secara sederhana, distribusi beta dapat diartikan sebagai distribusi probabilitas dari probabilitas. Dimana, beta distribusi ini mewakili semua nilai yang mungkin dari sebuah probabilitas ketika kita tidak tahu nilai probabilitasnya.

Hmmmm, apakah Anda sudah mulai bingung? Baiklah, bayangkan bahwa Anda akan bermain permainan mobile legend (atau permainan apapun yang Anda sukai). Ketika akan bermain, Anda akan berpikir atau bertanya: “Seberapa hebat saya dalam permainan ini?”. Misalkan “kehebatan” ini diukur dengan “**win rate**”, yang rumusnya adalah $\frac{\text{jumlah kemenangan}}{\text{total pertandingan}}$.

Misalkan, Anda bermain 3 pertandingan. Dalam 3 pertandingan ini, Anda menang 3 kali berturut-turut. Maka Anda akan memiliki metric win rate sebesar 3/3 atau 100%. Wah kalau begini, tentu Anda akan merasa bahwa Anda adalah orang terhebat dalam jagat permainan mobile legend karena Anda mendapatkan win rate sebesar 100% dalam 3 pertandingan pertama.

Lalu pertanyaannya, benarkah bahwa Anda sehebat itu dalam permainan ini? Faktanya, permainan awal dalam mobile legend ketika Anda membuat akun baru adalah melawan robot AI.

Sekarang, misalkan Anda sudah tidak melawan AI, Anda menjalani 3 pertandingan pertama melawan player lain. Kemudian dari 3 pertandingan tersebut, Anda mengalami kekalahan sebanyak 3 kali. Sehingga win rate Anda adalah 0/3 atau 0%.

Sekarang pertanyaannya adalah apakah benar Anda seburuk itu dalam permainan ini?

Jawaban untuk kedua pertanyaan diatas adalah **TIDAK**. Anda tidak sehebat itu dan Anda tidak seburuk itu pula. Lalu pertanyaannya seberapa baik Anda atau orang lain yang belum pernah memainkan permainan ini? Dalam kondisi inilah, distribusi beta diperlukan.

Kasus win rate yang kita hadapi dapat direpresentasikan oleh distribusi binomial (menang atau tidak dalam data pertandingan). Cara terbaik untuk merepresentasikan ekspektasi awal dari suatu kejadian adalah dengan menggunakan distribusi beta. Dijelaskan dalam distribusi beta bahwa kira-kira berapa win rate seseorang dalam sebuah permainan mobile legend sebelum dia memulai satu pertandingan pun.

Nilai distribusi beta berkisar antara 0 dan 1 yang mirip dengan range nilai probabilitas. Hal ini sudah cukup memberikan kesan bahwa distribusi beta dapat digunakan untuk memecahkan kasus yang kita hadapi. Seperti distribusi lainnya, beta distribusi juga memiliki property yang melekat. Property tersebut adalah *alpha* dan *beta*.

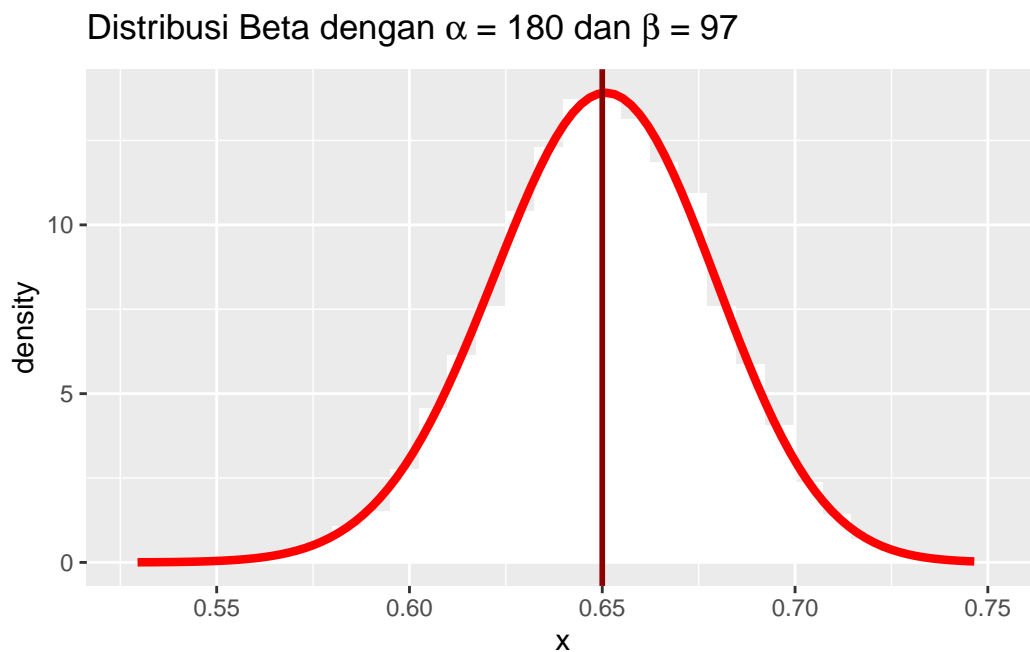
Misalkan kita mengetahui bahwa sepanjang satu season mobile legend, rata-rata tingkat kemenangan atau win rate seorang pemain adalah 65%, dengan nilai terendah-nya adalah 55% dan nilai tertinggi adalah 73%.

Dengan asumsi ini, kita bisa mengambil nilai $\alpha = 180$ dan $\beta = 97$. Ada 2 alasan kenapa saya memilih angka ini.

1. Karena nilai mean dari distribusi beta di mana $\alpha = 180$ dan $\beta = 97$ adalah $\sim 65\%$. Nilai mean dari distribusi didapat dengan formula $\frac{\alpha}{\alpha+\beta} = \frac{180}{180+97} = 65$.
2. Karena nilai dari distribusi beta dengan $\alpha = 180$ dan $\beta = 97$ akan berada di antara 55% dan 73%. Ingat bahwa distribusi beta harus menyimpan semua nilai yang mungkin muncul dari distribusi probabilitasnya.

Warning: `geom_vline()`: Ignoring ``mapping`` because ``xintercept`` was provided.

``stat_bin()`` using ``bins = 30``. Pick better value with ``binwidth``.



Seperti yang dibuktikan oleh gambar bahwa distribusi beta dengan $\alpha = 180$ dan $\beta = 97$, memiliki nilai tengah 65% dan berada di antara 55% dan 73%. Di sini, kita bisa katakan bahwa 65% akan menjadi angka performa awal bagi siapapun yang akan menjadi pemain mobile legend, termasuk Anda.

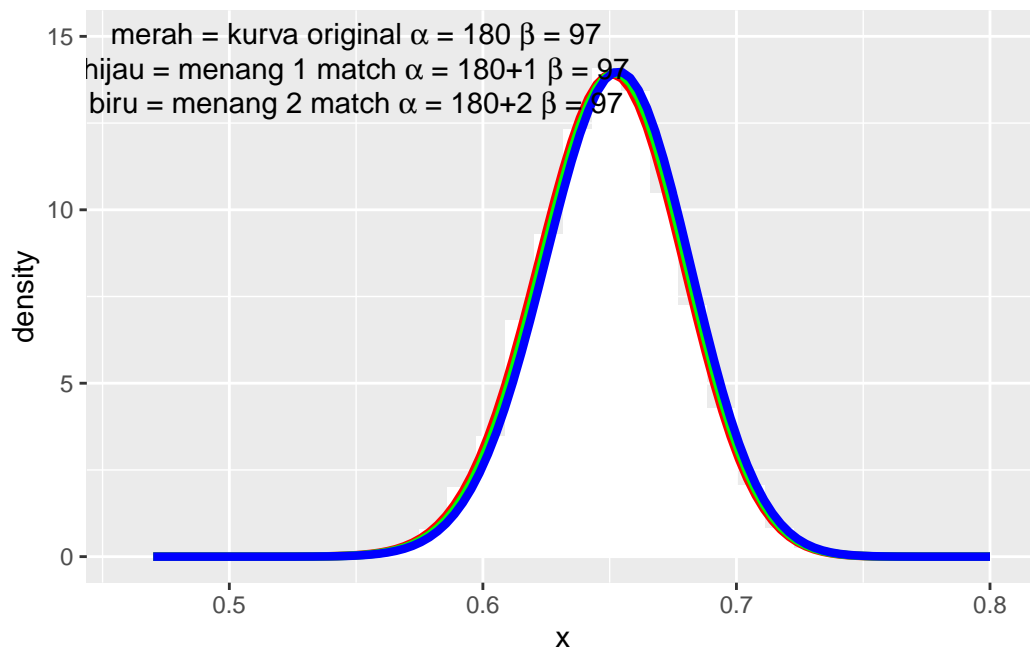
Jadi, ketika pertama kali bermain, performa “**win rate**” Anda adalah 65%, bukan 0% dan bukan pula 100%. Lalu bagaimana jika seandainya Anda memenangkan atau kalah dalam pertandingan? Ketika Anda memenangkan pertandingan, maka nilai alpha akan bertambah 1, dan jika Anda kalah dalam pertandingan, maka nilai beta akan bertambah 1.

```
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
Warning in is.na(x): is.na() applied to non-(list or vector) of type  
'expression'
```

```
Warning in is.na(x): is.na() applied to non-(list or vector) of type  
'expression'
```

```
Warning in is.na(x): is.na() applied to non-(list or vector) of type  
'expression'
```



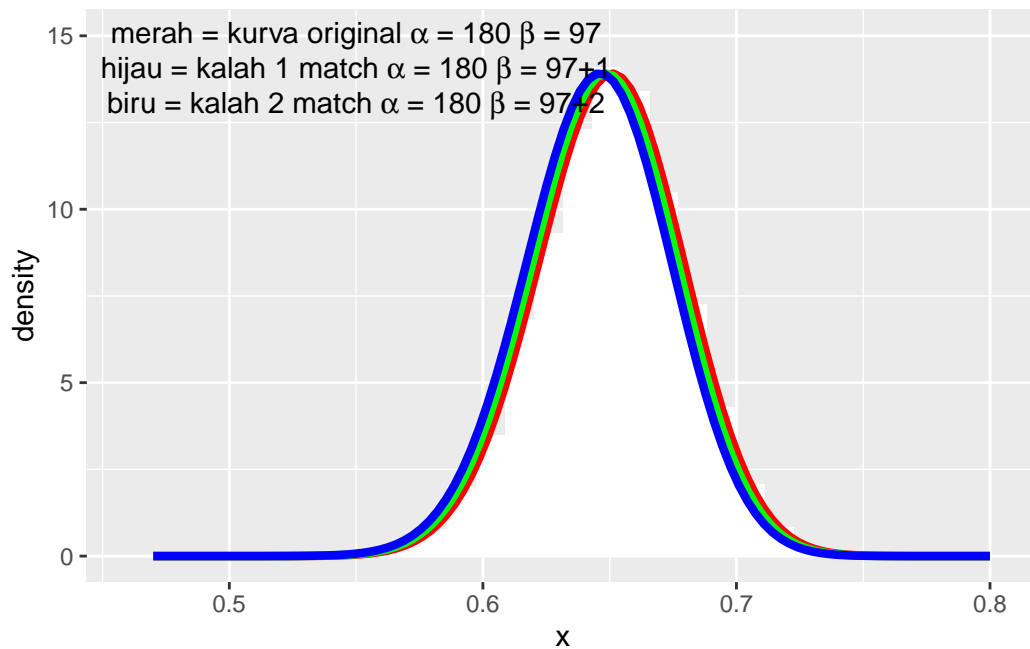
Nah, disini terlihat bahwa jika kamu hanya memenangkan pertandingan 2 kali berturut-turut (yang mungkin bisa saja hanya keberuntungan), performa **win rate** Anda tidak akan menjadi 100%, melainkan akan bertambah sedikit demi sedikit (bergerak dari 65% menuju ke 100%).

```
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

Warning in is.na(x): is.na() applied to non-(list or vector) of type
'expression'

Warning in is.na(x): is.na() applied to non-(list or vector) of type
'expression'

Warning in is.na(x): is.na() applied to non-(list or vector) of type
'expression'



Begitu juga sebaliknya, ketika Anda bertanding 2 kali berturut-turut dan mengalami kekalahan, performa **win rate** Anda tidak akan menjadi 0%, melainkan menurun sedikit demi sedikit dari 65% menuju 05.

Dari sini kita mengetahui bahwa distribusi beta adalah distribusi yang tepat untuk merepresentasikan distribusi probabilitas yang kita tidak tahu secara pasti nilainya. Namun kita masih memiliki data dan fakta yang cukup untuk membuat tebakan yang masuk akal.

Referensi : [Understanding the beta distribution \(using baseball statistics\)](#)