ISSN **2338-8102**

DESAIN ANIMASI STATISTIKA BERBASIS BAHASA PEMROGRAMAN R

I.G.A. Anom Yudistira
Universitas Sahid
Email: anom1392@gmail.com

Abstrak: Latar belakang penulisan makalah ini adalah bahwa, kebutuhan untuk mengembangkan animasi khususnya dalam Statistika sangat diperlukan terutama yang murah dan mudah dalam pengembangannya. Tulisan ini bertujuan untuk merancang suatu animasi untuk menjelaskan konsep/teori dalam statistika dengan memanfaatkan bahasa R yang merupakan *software open source*. Metode penelitian yang digunakan dalam penelitian ini adalah metode kepustakaan. Program animasi yang dirancang meliputi peragaan untuk konsep peluang dan fungsi sebaran peluang, konsep kuasa uji (*power test*) , konsep teorema limit pusat, dan konsep selang kepercayaan. Konsep-konsep lain dapat dibuat dengan teknik pemrograman yang serupa.

Kata kunci: statistika, animasi, bahasa R.

Abstract: The background of this paper is that, the need to inexpensive and easy to develop of animation in particular of Statistics is required. This paper aims to design an animation to explain the concept / theory in statistics using the R language which is open source software. The method used in this study is the literature studies of relevant subject. The program includes demonstration animations designed for the concept of probability distribution function, the concept of the power of the test (power test), the concept of central limit theorem, and the concept of confidence interval. Other concepts can be made with a similar programming technique.

Key words: statistics, animation, R programming

PENDAHULUAN

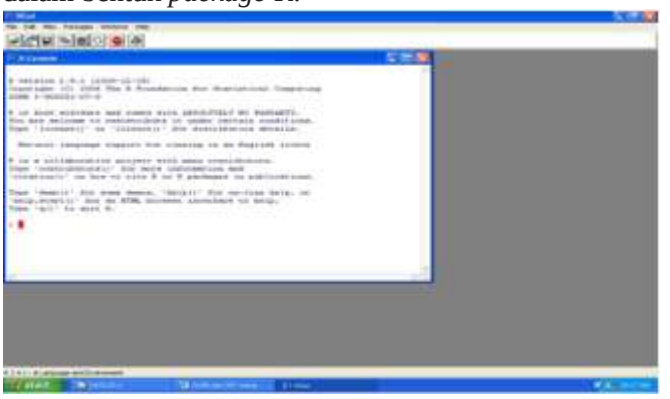
Latar belakang penulisan makalah ini adalah bahwa, kebutuhan untuk mengembangkan animasi khususnya dalam Statistika sangat diperlukan terutama yang murah dan mudah dalam pengembangannya. Alternatif yang dapat dipetimbangkan adalah menggunakan bahasa R yang merupakan versi sumber terbuka (*open-source*) dari bahasa pemrograman S (Alzola dan Harrel, 2006;4). Situs web resmi R tersedia di http://www.r-project.org. Versi komersial yang berbasis bahasa S adalah S plus. Menurut Venables dan Smith (2013; 2), bahasa R merupakan salah satu bahasa pemrograman yang memadukan kemampuan untuk memanipulasi data, kalkulasi dan peragaan graphik. Oleh karena itu bahasa R ini cukup baik dan memadai digunakan untuk melakukan simulasi model yang membutuhkan banyak peragaan graphik. Bahasa R memiliki kemampuan yang tidak kalah dangan paket-paket program pengolahan data komersial bahkan dalam beberapa hal kemampuannya lebih baik. Perbandingan R khususnya terhadap SAS dibahas secara rinci oleh Alzola dan Harrel (2006: 11). Bahasa R mendapat sambutan yang baik dari kalangan statistikawan di seluruh dunia, sayangnya di Indonesia belum banyak dikenal. Tulisan ini dibuat untuk mengenalkan kemampuan bahasa R khsususnya kalangan akademisi di Indonesia, dalam hal membantu dalam pengajaran statisika, dan disamping itu untuk mendukung gerakan IGOS (*Indonesia Go Open Source*).

Permasalahan dalam tulisan ini meliputi: (1) Bagaimanakah mengembangkan program bahasa R untuk animasi? dan (2) Bagaimanakah mengembangkan kemampuan animasi bahasa R untuk diterapkan dalam statistika? Tulisan ini bertujuan untuk: (1) memberikan gambaran mengenai kemampuan animasi bahasa R dan pemanfaatannya untuk demo dalam pengajaran statistika; (2) mengeksplorasi kemampuan bahasa R khususnya untuk animasi dalam memperaga-kan suatu konsep/teori statistika. Versi yang digunakan adalah R 2.15.3 di bawah sistem operasi Microsoft Windows XP. Kemampuan komputasinya dicoba secara langsung, mengunakan Laptop Intel Core 2 Duo CPU T5750 dan 0.99 GB RAM. Metode yang digunakan dalam tulisan ini adalah metode kepustakaan, dengan pendekatatan deskriptif dan eksploratif.

PEMBAHASAN

Cara Kerja Bahasa R

R merupakan sebuah paket dan sekaligus bahasa pemrograman untuk analisis data dan grafik, dan merupakan bahasa pemrograman tingkat tinggi (*hight level programming*). R dapat digunakan secara interaktif sehingga hasil perhitungan segera dapat dilihat, tetapi apabila perhitungannya kompleks maka perintah-perintah R ditulis lebih dahulu dalam *text editor*, kemudian dipanggil dengan fungsi source. Program R dapat diunduh gratis di internet di http://cran.r-project.org/, dan sebagaimana halnya program *open source*, R dirancang secara *voluntary* oleh ahli-ahli statistik dan pemrograman di seluruh dunia, sehingga perkembangannya sangat pesat, Versi R terakhir ketika tulisan ini dibuat adalah 2.15.3. R juga dilengkapi dengan program-program kecil tambahan (*add-in*) yang disebut dengan *package*. *Package* memberikan kemampuan tambahan, misalnya perhitungan teknik-teknik statistik yang canggih, *interface* dan lain-lain. *Packages* ini juga dapat diunduh gratis. Secara periodik muncul *package- package* baru yang menarik yang dapat dipilih sesuai kebutuhan. Semua *package* ini dapat diperoleh gratis di http://cran.r-project.org/. Gambar 1 berikut ini memberikan GUI (*Graphics User Interface*) / antar muka grafis standard dari bahasa R. Terdapat banyak GUI lain yang juga dapat diperoleh gratis antara lain SciViews, RStudio dan Rcmdr. Rcmdr dikemas dalam bentuk *package* R.



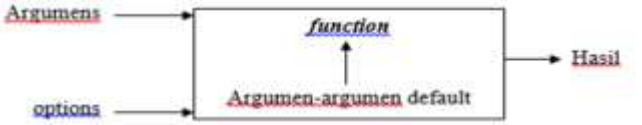
Gambar 1. Antar Muka Grafis Bahasa R

Bahasa R dan dengan demikian juga S merupakan bahasa yang dirancang untuk komputasi statistik, grafik dan matriks.

Verzani (2004; 1) mengemukakan bahwa keuntungan-keuntungan yang diperoleh bila menggunakan bahasa R untuk pengolahan data dan statistik antara lain: R disamping dapat diunduh gratis dan juga dapat dijalankan pada berbagai sistem operasi (UNIX, Windows, Macintosh), sintaksnya mudah dipelajari dan memiliki banyak sekali fungsi-fungsi statistik terpasang. Disamping keuntungan tersebut Verzani (2004;1) juga mengemukakan kelemahan utama bahasa R adalah tidak adanya dukungan komersial. Alzola dan Harrel (2006; 11) memberikan perbandingan rinci atas kemampuan bahasa R terhadap program komersial SAS. S-Plus sebagai "kembaran" komersial dari bahasa R memiliki antar muka yang lebih baik, tetapi berjalan sedikit labih lambat dari pada R.

Skema Kerja Bahasa R

Setelah R terpasang pada komputer maka dapat diakses melalui *shortcut* atau menu *start*. Prompt default ">" mengindikasikan bahwa R menunggu perintah yang diberikan, baik berupa fungsi maupun objek lain. R merupakan suatu bahasa berorientasi objek, artinya bahwa variabel, data, fungsi, hasil dan sebagainya, disimpan dalam memori aktif komputer dalam bentuk objek dan mempunyai sebuah nama. Pengguna dapat mengenakan aksi terhadap objek-objek tersebut melalui suatu operator (aritmatika, logika, dan perbandingan) dan fungsi (*functions*). Fungsi itu sendiri juga merupakan suatu objek. Penamaan terhadap objek bersifat *case sensitive* (membedakan huruf besar dan huruf kecil). Jadi X dan x merupakan objek yang berbeda. Fungsi R dapat disketsakan sebagai berikut.



Gambar 2. Sketsa fungsi R.
Sumber: Paradis (2005; 5)

Argumen suatu fungsi dalam R yang juga merupakan suatu objek ("data", formula, expresi, dan sebagainya) merupakan input dari fungsi tersebut. Beberapa fungsi menyediakan nilai default dari argumennya, nilai ini dapat diubah oleh penggunanya dengan *options* tertentu. Beberapa fungsi tidak membutuhkan nilai argumen dalam menjalankan perintahnya, ini dapat terjadi karena ada dua

kemungkinan: (1) Semua argumennya mempunyai nilai default (karenanya dapat diubah dengan *options*). Jadi dalam mengeksekusi fungsi tersebut tidak perlu memasukkan nilai argumen, karena fungsi akan menggunakan nilai default sebagai nilai argumen. (2) Tidak ada argumen yang didefinisikan sebelumnya pada fungsi tersebut.

Objek-objek R (termasuk fungsi) dikemas dalam bentuk *add-in* (perangkat lunak tambahan) yang oleh R disebut dengan *package*. Pada saat R dipanggil pertama kali, maka ada tujuh *package* yang dipanggil dan disimpan dalam memori aktif. Setiap *package* memberikan kemampuan tambahan untuk R.

Bekerja dengan Data

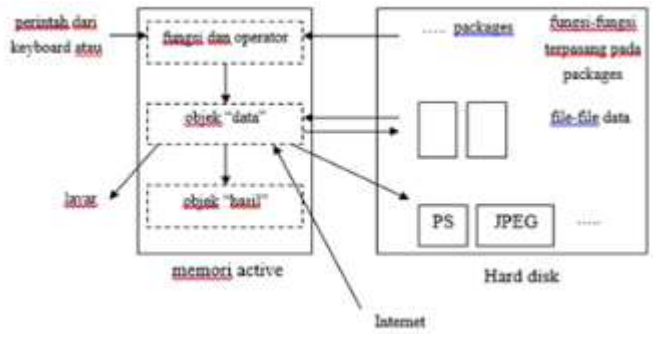
R bekerja berdasarkan jenis struktur datanya. Struktur data paling sederhana adalah vector. Vektor dibuat dengan fungsi c, hasilnya disimpan dalam suatu objek. Berikut ini adalah teladan tipe vektor numerik.

```
> x <- c(10, 4.5, 6, 7, 2.3, 7.1)
```

Objek x menyimpan nilai numerik 10, 4.5, 6, 7, 2.3, dan 7.1. Perintah untuk menampilkan isi dari objek x adalah.

```
> x
[1] 10.0  4.5  6.0  7.0  2.3  7.1
```

Gambar berikut ini memperagakan skema kerja bahasa R:



Gambar 3. Skema kerja bahasa R.
Sumber: Paradis (2005; 4).

Peralatan Animasi Pada
Penggunaan Fungsi Sys.sleep()

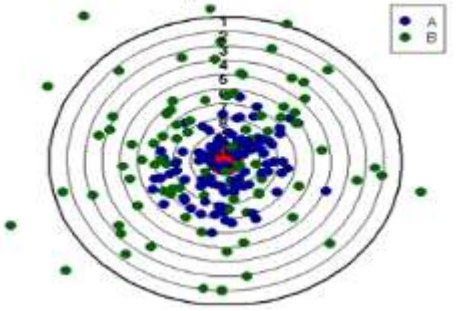
Kebanyakan programmer akan berpikir bahwa GIF *images* merupakan pilihan yang baik untuk membuat animasi grafis, karena GIF sudah dikenal baik sebagai salah satu dari sedikit format gambar yang digunakan untuk membuat animasi. Bagi pengguna R, animasi dapat dibuat dengan menggunakan fungsi terpasang Sys.sleep() dalam sebuah *loop*. Fungsi ini bertujuan untuk memperlambat *loop*, sehingga pengguna dapat dengan jelas memperlihat proses secara perlahan dalam loop. Yihui Xie (2008; 6), membuat peragaan animasi untuk merotasi kata "Animation", seperti dalam *loop* di bawah ini.

```
for (i in 1:360) {
plot(1, ann = F, type = "n", axes = F)
text(1, 1, "Animation", srt = i, col = rainbow(360)[i],
 cex = 7 * i/360)
Sys.sleep(0.01)
}
```

Animasi berikut ini memberikan teladan untuk memnggambarkan perbedaan variasi dua data, yang dianalogikan dengan penyebaran sasaran tembak. Semakin besar variasi suatu data, semakin menyebar sasaran tembaknya.

```
a <- seq(0,2*pi, length.out=100)
plot(cos(a),sin(a),type="l",lwd=2,xlab="",ylab="",
axes=FALSE, asp=1)
for (i in seq( 0.1,0.9,by=0.1)) lines(i*sin(a), i*cos(a))
polygon(sin(a)*0.1,cos(a)*0.1,col="red")
for (i in 1:10){
text(x=0,y=i/10-0.025,labels=11-i,font=2)
shots1 <- matrix(ncol=2, data=rnorm(n=20,sd=0.2))
shots2 <- matrix(ncol=2, data=rnorm(n=20,sd=0.5))
points(shots1,col="darkblue",pch=16,cex=1.5)
points(shots2,col="darkgreen",pch=16,cex=1.5)
Sys.sleep(1)
}
legend("topright",legend=c("A","B"),pch=16,
col=c("darkblue","darkgreen"),pt.cex=1.5)
```

Fungsi Sys.sleep() mempunyai satu argumen yaitu time. Argumen *time* ini memberikan nilai interval waktu dalam suatu *loop*. Jadi semakin besar nilai time semakin lama animasi tersebut berjalan.



Gambar 4 Perbedaan penyebaraan sasaran untuk dua data dengan variasi berbeda.

Gambar di atas akan bergerak secara berlahan-lahan untuk memplot titik-titik berwarna hijau dan biru. Sasaran hijau (titik-titik hijau) merepresentasikan data dengan nilai

penyebaran yang lebih besar (simpangan baku 0.5). Sasaran warna biru (titik-titik biru) merepresentasikan data dengan penyebaran yang lebih kecil (simpangan baku 0.2).

Package **animation**

R menyediakan sebuah *package* khusus untuk animasi, yang dipanggil dengan perintah library(animation). *Package* ini ditulis oleh Yihui Xie dan *Package* ini dibuat utamanya untuk penggunaan HTML dan JavaScript serta dapat pula untuk grafik window dari R (yaitu x11) untuk peragaan statistik. Output lain dalam bentuk Flash (SWF) dan GIF / MPG dapat dihasilkan, asalkan *software* yang diperlukan sudah terinstal. Versi teakhir dari *package* ini adalah 2.2, dapat di unduh gratis pada cran.r-project.org.

Package ini memberikan 33 buah fungsi tambahan untuk R. Fungsi tersebut salah satunya adalah untuk memperagakan sample.cluster(), yaitu untuk memperagakan penarikan sample acak klaster. Perintah R berikut ini memanfaatkan fungsi sample.cluster untuk memperagakan pengacakan pada penarikan sample acak berklaster. Keluaran dari perintah ini merupakan halaman HTML.

```
oopt = ani.options(interval = 1, nmax = 30)
ani.options(ani.height = 350, ani.width = 500,
outdir = getwd(), nmax = 30, interval = 1,
title = "Peragaan cluster sampling", description = "Pada setiap cluster yang terpilih, semua elemen-epemenya akan dipilih sebagai sample.")
ani.start()
par(mar = rep(1, 4), lwd = 2)
sample.cluster(col = c("bisque", "white"))
ani.stop()
ani.options(oopt)
```

Package **rpanel**

Package ini memberikan sekumpulan fungsi (ada sedikitnya 28 fungsi) untuk mengembangkan kontrol antar muka grafis (GUI) yang sederhana. Fungsi-fungsi ini dibangun dengan memanfaatkan *package* tcltk, sehingga memudahkan programmer membangun GUI, untuk membuat suatu animasi dibanding langsung memanfaatkan *package* tcltk. Pemanfaatan fungsi-fungsi dalam *package* rpanel, memungkinkan *user* merubah nilai-nilai parameter pada saat menjalankan animasi, dalam bentuk *slider*, *double button* dan banyak lagi panel-panel kontrol yang lebih canggih. Teladan berikut ini menampilkan animasi grafik fungsi kepekatan normal dengan merubah nilai-nilai tingkat kemulusannya. Pada fungsi density dengan argumen bw memberikan nilai *bandwidth* pemulusannya.

```
density.draw <- function(panel) {
plot(density(panel$x, bw = panel$h))
        panel
}
panel <- rp.control(x = rnorm(50))
rp.slider(panel, h, 0.5, 5, log = TRUE, action = density.draw)
```

Perintah di atas akan menggambarkan animasi grafik fungsi kepekatan normal baku, yang datanya diperleh dari dari nilai x. *Slider* akan mengontrol peubah h dengan batas bawah 0.5 dan batas atas 5. Gambar di bawah ini memberikan bentuk *slider* yang diberikan.

Gambar 5. Contoh slider menggunakan fungsi rp.slider.

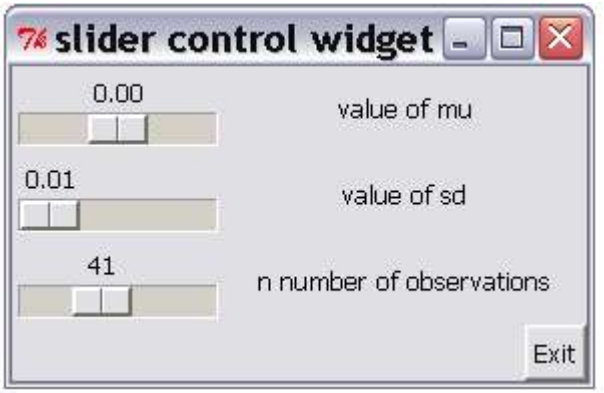
Nilai variabel h akan diubah-ubah dengan menjalankan *slider*-nya, nilai h ini akan diteruskan ke argumen bw pada fungsi density.

Package **TeachingDemos**

Package ini mengandung fungsi-fungsi untuk demo dalam pengajaran statistik di dalam kelas. Pada *package* ini sedikitnya terdapat 60 fungsi yang dapat dimanfaatkan untuk mengembangkan animasi, yang didesain untuk keperluan pengajaran statistik tertentu. Program berikut ini memperagakan *scatter plot* dari sebuah sampel yang berasalah dari berbagai populasi normal, dengan berbagai parameter μ dan α, serta berbagai ukuran sampel (n). Nilai-nilai μ,α, dan n diubah-ubah melalui sebuah *slider*.

```
plot.sample.norm<-function(){
refresh.code<-function(...){
  mu<-slider(no=1); sd<-slider(no=2); n<-slider(no=3)
 x<-rnorm(n,mu,sd)
plot(x)
}
slider(refresh.code,sl.names=c("value of mu","value of sd",
     "n number of observations"),sl.mins=c(-10,.01,5),
     sl.maxs=c(+10,50,100), sl.deltas=c(.01,.01,1),
     sl.defaults=c(0,1,20))
}
plot.sample.norm()
```

Penggunaan fungsi *slider* dari *package* TeachingDemos lebih intuitif diban-dingkan dengan fungsi rp.slider. Keluaran dari program tersebut adalah:



Gambar 6. Contoh slider memanfaatkan fungsi slider dari package TeachingDemos

Animasi Menggunakan R

Animasi Sebaran Peluang Normal

Fungsi yang diciptakan untuk animasi ini diberi nama run.ProbNormal(). Fungsi ini memanfaatkan *package* TeachingDemos dan tcltk. *Slider* yang ditampilkan seperti berikut ini:

Gambar 7. Kontrol Panel dihasilkan oleh run.ProbNormal().

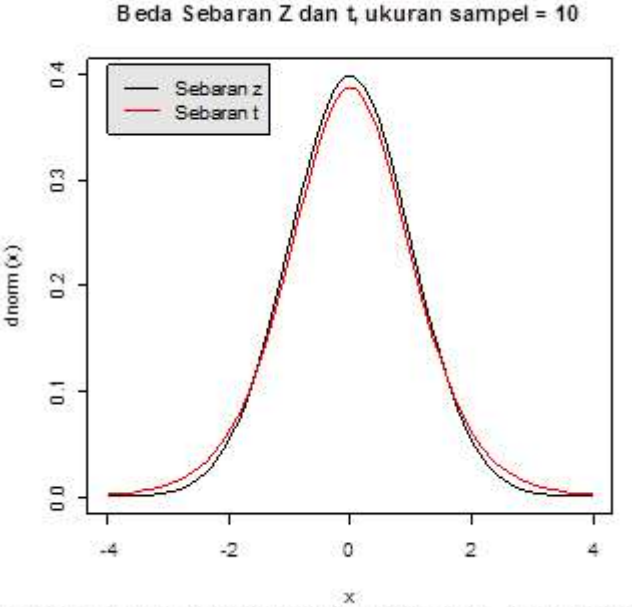
Animasi yang dihasilkan oleh fungsi run.Probnormal(), adalah memperagakan nilai peluang (luas daerah di bawah kurva) dari suatu sebaran normal, dengan berbagai rata-rata dan / atau simpangan baku, serta batas atas dan / batas bawah. Keempat parameter tersebut dapat diganti-ganti dengan menggeser-geser *slider* yang bersesuaian.

Animasi perbedaan kurva normal baku dan t-student

Animasi ini bertujuan untuk membandingkan kurva normal baku dengan kurva t-student pada berbagai ukuran sampel. Besarnya ukuran sampel diganti-ganti dengan mengeser-geser *slider*. Mahasiswa diharapkan memiliki pemahaman bahwa sebaran t dan sebaran normal baku akan semakin mirip (dekat) dengan semakin bertambahnya ukuran sampel. Jadi apabila ukuran sampel yang diambil besar (lebih dari 30), maka penggunaan tabel Z dan t menjadi tidak berbeda. Gambar di bawah ini memperlihatkan *slider* dan grafik yang bersesuaian.

Gambar 8. Kontrol Panel dihasilkan oleh run.beda.kurva().

Grafik yang dihasilkan oleh program animasi fungsi run.beda.kurva() diperagakan oleh Gambar berikut ini:

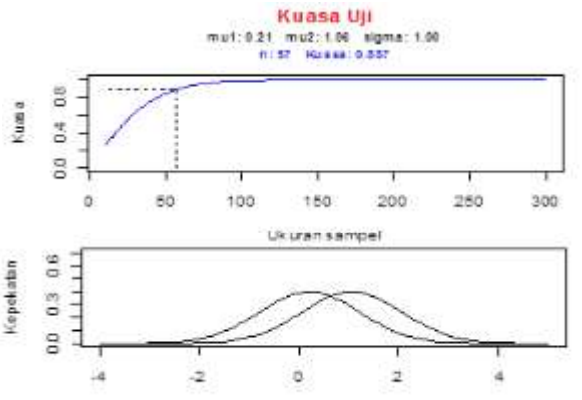


Gambar 9. Grafik perbedaan kurva normal baku dan t-student.

Animasi untuk kuasa uji.

Animasi ini ditampilkan dengan mengetik perintah rp.kuasa() pada *console* R. Fungsi rp.kuasa akan memberikan animasi untuk nilai kuasa uji, apabila ukuran sampel (n) berubah dan / atau nilai rata-rata yang diasumsikan dalam H0 (mu1) dan / atau nilai rata-rata sebenarnya (mu2) juga berubah. Kontrol panel yang dihasilkan adalah seperti di bawah ini.

Gambar 10. Kontrol Panel dihasilkan oleh rp.kuasa(). Grafik berikut ini merupakan gambar yang dianimasi oleh kontrol panel.



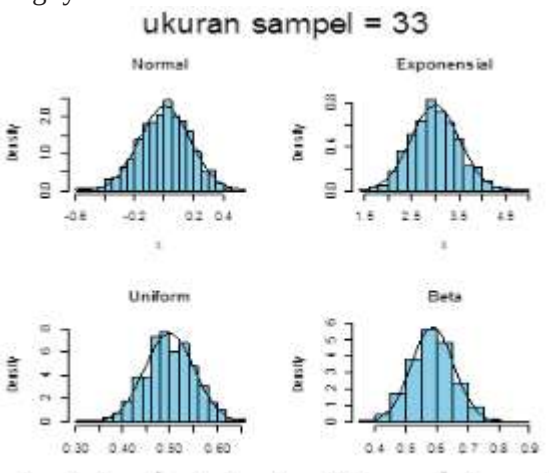
Gambar 11. Grafik yang dianimasi oleh fungsi rp.kuasa().

Animasi Teorema Limit Pusat

Animasi ini memperagakan bentuk sebaran sampling (), yang nilai individualnya berasal dari empat sebaran utama yakni Normal, Eksponensial, Uniform, dan Beta. Nilai default ukuran sampelnya adalah 1, dan ulangan sampelnya dilakukan 1000 kali. Jadi nilai default ini sama dengan sebaran individual (*X*). Kontrol panel fungsi ini adalah:

Gambar 12. Kontrol panel berupa slider yang dihasilkan oleh fungsi run.limit.pusat().

Tujuan dari animasi ini adalah untuk memperagakan teorema limit pusat, dengan menambahkan ukuran sampel, ternyata apapun sebaran asal, maka bentuk sebaran samplingnya akan mendekati bentuk sebaran normal.



Gambar 13. Animasi grafik fungsi run.limit.pusat() dengan *slider* ukuran sampel digerakan ke 33 (ukuran sampel 33)

Animasi Selang Kepercayaan

Animasi ini memperagakan 100 selang kepercayaan berbasiskan sebaran z atau t. Tingkat kepercayaan dapat diubah-ubah pada *slider*, dan akan terlihat beberapa selang kepercayaan yang tidak mencakup nilai rata-rata sebenarnya. Tujuan animasi ini adalah memberikan pemahaman pada perserta didik mengenai konsep selang kepercayaan. Fungsi R untuk animasi ini diberi nama run.sk().

PENUTUP

Kesimpulan

1. Program animasi dapat dikembangkan berbasis bahasa R dengan memenfaatkan fungsi Sys.sleep(), *Package* animation, rpanel dan TeachingDemos. Fungsi dan ketiga *package* tersebut cukup efektif dalam membuat animasi.
2. Pemanfaatan fungsi dan *packages* yang disediakan oleh R tersebut akan memudahkan dalam membuat program animasi untuk memperagakan suatu teori dalam statistika.

Saran-saran

1. Pada tahap berikutnya akan sangat manarik bila dikembangkan animasi statistik berbasis bahasa R, yang dapat membantu mahasiswa dalam meningkatkan pemahaman atas suatu teori statistika, dan diukur seberapa jauh efektifitasnya.
2. Perlu juga dikembangkan suatu *package* R yang sesuai kurikulum mata kuliah statistika pada suatu program pendidikan tertentu dan digunakan sebagai modul pembelajaran / praktikum.

DAFTAR PUSTAKA

Alzola, C., and F. Harrell. 2006. An Introduction to S and The Hmisc and Design Libraries. http://biostat.mc.vanderbilt.edu/RS., diakses tanggal 8 Januari 2007.

Paradis, E. 2005. R for Beginners. Institut des Sciences de l'Evolution, Univerisite Montpellier II. France. http://cran.r-project.org/., diakses tanggal 7 Desember 2005.

Venables, W.N., D.M. Smith, and the R Development Core Team. 2013. An Introduction to R. http://www.r-project.org/., diakses tanggal 7 Desember 2013.

Versani, J. (2002). simpleR – Using R for Introductory Statistics. .math.csi.cunv.edu/Statistics/R/simpleR/Simple., diakses tanggal 7 Desember 2009.

Yihui Xie. 2008. Package animation: Animated Statistics Using R. http://cran.r-project.org/., diakses tanggal 12 Oktober 2012.