

**METODE – METODE PENGUJIAN UNTUK HIPOTESIS
BERGANDA**

INTAN PERMATA SARI

0304010293



**UNIVERSITAS INDONESIA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
DEPARTEMEN MATEMATIKA
DEPOK
2009**

**METODE – METODE PENGUJIAN UNTUK HIPOTESIS
BERGANDA**

**Skripsi diajukan sebagai salah satu syarat untuk memperoleh
gelar Sarjana Sains**

Oleh:

INTAN PERMATA SARI

0304010293



DEPOK

2009

SKRIPSI : METODE-METODE PENGUJIAN UNTUK HIPOTESIS
BERGANDA
NAMA : INTAN PERMATA SARI
NPM : 0304010293

SKRIPSI INI TELAH DIPERIKSA DAN DISETUJUI
DEPOK, 16 DESEMBER 2009

DRA. SASKYA MARY, M.Si

PEMBIMBING I

DRA. SITI NURROHMAH, M.Si

PEMBIMBING II

Tanggal lulus Ujian Sidang Sarjana : 21 Desember 2009

Penguji I : Dra. Saskya Mary, M.Si

Penguji II : Dra. Rianti Setiadi, M.Si

Penguji III : Dra. Yahma Wisnani, M.Kom

KATA PENGANTAR

Alhamdulillah rabbi 'aalamiin. Segala puji dan syukur hanya kepada ALLAH SWT, yang telah memberikan banyak cobaan serta melimpahkan rahmat dan karunia-Nya kepada penulis sehingga skripsi ini dapat diselesaikan dengan baik, walaupun dengan tertatih-tatih. Shalawat dan salam penulis sampaikan kepada suri tauladan kita, manusia biasa dengan akhlak luar biasa, Rasulullah SAW, dengan semangat dan perjuangan beliau-lah yang membuat penulis bisa bertahan hingga detik ini. terselesaikannya skripsi ini tidak terlepas dari bantuan, bimbingan, dorongan, dan doa yang tulus dari banyak pihak. Oleh karena itu, pada kesempatan ini, penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Ibu tercinta, yang selama ini tak pernah berhenti untuk selalu mendukung dan mendoakan penulis, semoga lekas sembuh dan tetap kuat menjalani pengobatan kemoterapi & sinarnya. Bapak tercinta, yang selama ini tak pernah mengeluh dan terus bekerja keras walaupun cobaan selalu menghampiri keluarga kami. Dan adik tersayang, semoga bisa segera menyelesaikan kuliahnya. *“Ya ALLAH...., berilah kekuatan dan kesabaran kepada kami agar selalu tetap istiqomah di jalan-Mu...”*
2. Ibu Saskya Mary selaku Pembimbing 1 penulis yang telah meluangkan waktunya untuk memberikan bimbingan, saran, pengarahan dan kemudahan lainnya dengan sangat sabar sehingga skripsi ini dapat diselesaikan dengan baik.

3. Ibu Siti Nurrohmah selaku Pembimbing 2, yang juga telah meluangkan waktu disela kesibukkan. Terima kasih banyak atas saran dan pengarahan yang ibu berikan selama penulisan skripsi ini.
4. Lek husni yang sudah bersedia meminjamkan laptop, membantu memperbaiki komputer dan salma sayang, adik kecilku dengan celotehannya bisa menghiburku di sela-sela kepenatan.
5. Mbah kakung, Mbah Putri, Pak De, Bu De, Pak Lek, Bu Lek, saudara-saudara sepupu dan seluruh keluarga besar penulis yang banyak memberikan dukungan dan doa.
6. Ibu Siti Aminah dan Ibu Yekti Widyaningsih selaku pembimbing akademis yang telah memberikan nasihat dan bimbingannya.
7. Seluruh dosen Departemen Matematika atas segala ilmu yang penulis peroleh selama menjadi mahasiswa Matematika UI
8. Seluruh karyawan Departemen Matematika yang telah banyak memberikan bantuannya.
9. Avi, Desti, Nurma, Mia, Rani, Raisa, Miranti, Nabung, Teman-teman seperjuangan penulis yang sama-sama berjuang untuk menyelesaikan skripsi pada semester ini. Sebentar lagi kita bebas..... ALLAHU AKBAR !!!!
10. Spina dan Leli, tetap semangat ya... Allah pasti punya rencana yang indah untuk kalian...
11. Lisa, Nola, Dina, las, lif, atas support, bantuan dan doanya yang selalu menyempatkan diri bertanya kepada penulis: "How are you?"

12. Teman-teman angkatan 2004, atas persahabatan kita selama ini, dukungan, bantuan, support dan doa, semoga silaturahmi kita terus terjalin.
13. Teman-teman angkatan 2001, 2002, 2003, 2005, 2006, 2007
14. Teman-teman Ligo, beladiri Thifan, tahsin Utsmani, renang GTA, alumni SMUN 14, yang telah mengisi hari-hari penulis, berbagi ilmu, sharing dan pengalaman hidup.
15. Adik-adik binaan Matematika UI 2006, SMUN 14 angk2008, SMUN 14 kelas XI, adik-adik murid privat dan bimbel yang telah memberikan inspirasi kepada penulis untuk terus bersemangat.
16. Abu Bakar, Umar, Mush'ab, Lintang, Ikal, Arai, Delisa, Janggem, Deokman, Hwarang Kim Yu Shin, Conan, Harry Potter, Kobe Bryant, Tung Fang Xiang, Kick Andy, Ustadz, Ustadzah, dr.Walta, dr.Hilman, suster Maya, suster Maria, pasien di RS Dharmais, tokoh-tokoh dan orang-orang yang telah menginspirasi penulis untuk terus berjuang meraih mimpi-mimpi, selalu optimis, dan tak kan pernah menyerah.
17. Semua pihak yang telah membantu penulis dengan dukungan dan doanya. Semoga skripsi ini dapat berguna bagi siapa saja yang mengkajinya, serta dapat dikembangkan dan disempurnakan agar lebih bermanfaat untuk kepentingan orang banyak.

Penulis

2009

ABSTRAK

Dalam pengujian hipotesis berganda, dilakukan pengujian lebih dari satu hipotesis, yang diuji pada satu waktu secara simultan. Apabila masing-masing pengujian dalam suatu *family hipotesis* mempunyai probabilitas melakukan kesalahan tipe 1, maka secara keseluruhan pada pengujian hipotesis berganda akan terjadi penggandaan probabilitas kesalahan tipe 1. Probabilitas melakukan kesalahan tipe 1 pada pengujian hipotesis berganda akan semakin membesar seiring dengan meningkatnya jumlah pengujian. Untuk mengatasi hal itu, ada beberapa cara untuk mengukur kesalahan tipe 1 dalam *family hipotesis* diantaranya *Family Wise Error Rate* (FWER), *False Discovery Rate* (FDR), dan *positif False Discovery Rate* (pFDR). Untuk mengontrol kesalahan tersebut, diperlukan suatu metode sedemikian sehingga probabilitas kesalahan tipe 1 keseluruhan $\leq \alpha$. Pada tugas akhir ini, akan dibahas metode - metode pengujian untuk hipotesis berganda yaitu metode Bonferroni yang merupakan salah satu metode untuk FWER, metode Benjamin-Hochberg untuk FDR yang memperbaiki Metode Bonferroni dan metode Storey untuk pFDR yang memperbaiki Metode Benjamin-Hochberg.

Kata kunci : *family hipotesis*, kesalahan tipe 1, pengujian hipotesis berganda
ix + 65 hlm.; lamp

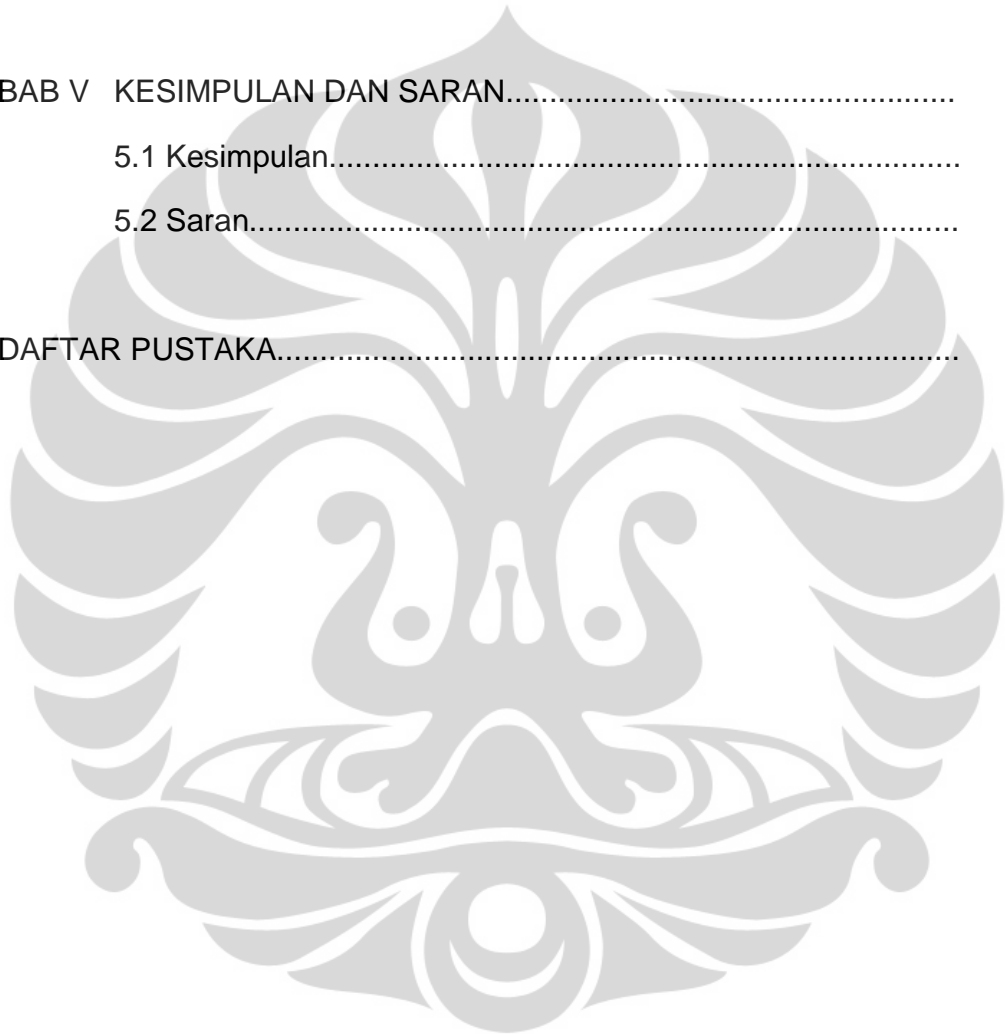
Bibliografi: 7 (1995-2003)

DAFTAR ISI

	Halaman
KATA PENGANTAR.....	i
ABSTRAK.....	iv
DAFTAR ISI.....	v
DAFTAR TABEL.....	viii
DAFTAR LAMPIRAN.....	ix
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Perumusan Masalah.....	2
1.3 Tujuan Penulisan.....	3
1.4 Pembatasan Masalah.....	3
1.5 Sistematika Penulisan.....	3
BAB II LANDASAN TEORI.....	5
2.1 Variabel Random.....	5
2.2 Ekspektasi.....	8
2.2.1 Ekspektasi Bersyarat.....	8
2.3 Sampel Random.....	10
2.4 Distribusi Statistik Terurut.....	10
2.5 Hukum De Morgan.....	12
2.6 Teorema Bayes.....	14

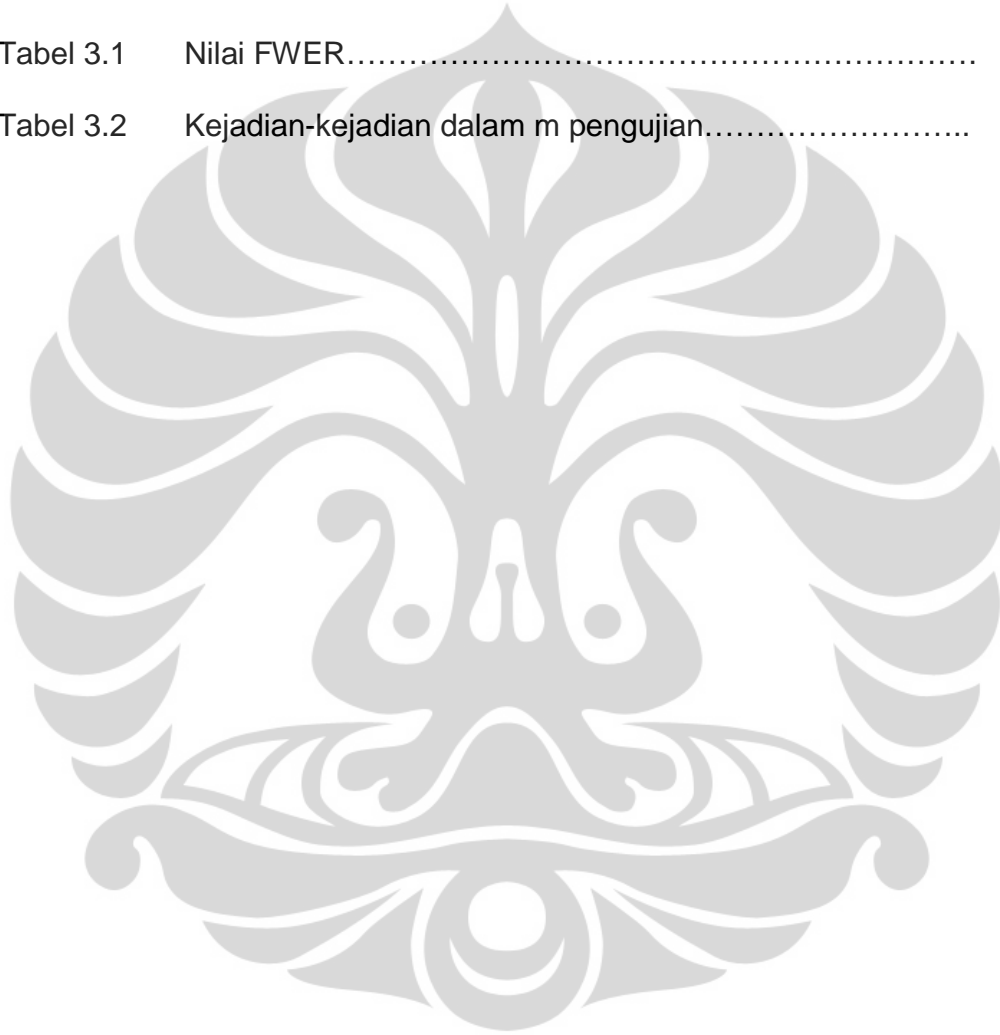
2.6.1 Probabilitas dan Partisi.....	14
2.6.2 Teorema Bayes.....	15
2.7 Pengujian Hipotesis.....	15
2.7.1 Hipotesis Statistik.....	16
2.7.2 Statistik Uji.....	19
2.7.3 Aturan Keputusan.....	21
2.8 Pengujian Hipotesis Tunggal.....	21
BAB III METODE-METODE PENGUJIAN UNTUK HIPOTESIS BERGANDA.....	24
3.1 Pengertian Pengujian Hipotesis Berganda.....	24
3.2 <i>Family Wise Error Rate</i> (FWER).....	26
3.2.1 Metode Bonferroni.....	27
3.3 <i>False Discovery Rate</i> (FDR).....	29
3.3.1 Perbandingan Antara FWER dan FDR.....	32
3.3.2 Metode Benjamin dan Hochberg.....	33
3.4 <i>Positif False Discovery Rate</i> (pFDR).....	41
3.4.1 Perbandingan Antara FDR dan pFDR.....	42
3.4.2 Metode Storey.....	42
3.4.3 Interpretasi Bayesian dari pFDR.....	43
3.4.4 Menaksir pFDR.....	47
3.4.5 <i>q-value</i>	48
BAB IV APLIKASI PENGUJIAN HIPOTESIS BERGANDA PADA DATA DNA MICROARRAY EKSPERIMENT.....	49

4.1 Latar Belakang Masalah.....	49
4.2 Permasalahan.....	51
4.3 Sumber Data.....	51
4.4 Analisis Data.....	52
BAB V KESIMPULAN DAN SARAN.....	56
5.1 Kesimpulan.....	56
5.2 Saran.....	57
DAFTAR PUSTAKA.....	58



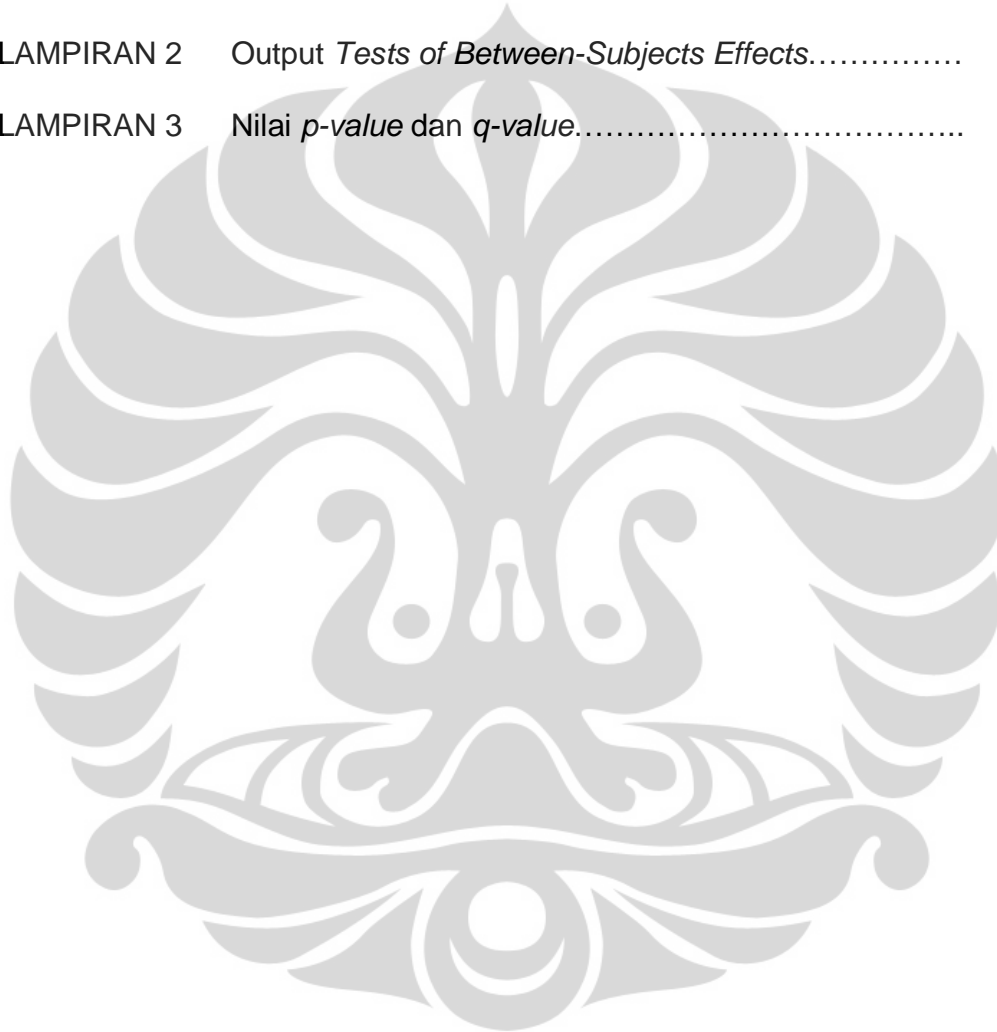
DAFTAR TABEL

	Halaman
Tabel 2.1 Probabilitas kejadian dari pengujian tunggal.....	22
Tabel 3.1 Nilai FWER.....	27
Tabel 3.2 Kejadian-kejadian dalam m pengujian.....	30



DAFTAR LAMPIRAN

	Halaman
LAMPIRAN 1 Output <i>Multivariate Test</i>	59
LAMPIRAN 2 Output <i>Tests of Between-Subjects Effects</i>	60
LAMPIRAN 3 Nilai <i>p-value</i> dan <i>q-value</i>	63



BAB I

PENDAHULUAN

1.1 LATAR BELAKANG

Pengujian merupakan salah satu bagian yang penting dalam statistika inferensi. Statistika inferensi adalah metode-metode analisis terhadap sampel untuk mendapatkan pendugaan atau penarikan kesimpulan tentang keseluruhan informasi dalam populasi. Pengujian dilakukan terhadap parameter populasi dalam bentuk hipotesis statistik, yaitu hipotesis null dan hipotesis alternatif.

Pengujian hipotesis adalah suatu aturan yang mengacu pada keputusan menerima atau menolak hipotesis statistik yang diuji. Keputusan menolak atau tidak menolak hipotesis biasanya dilakukan terhadap hipotesis null. Dalam pengujian hipotesis, terdapat dua jenis kesalahan yang mungkin terjadi. Kesalahan tipe 1 terjadi apabila hipotesis null ditolak ketika hipotesis ini benar. Kesalahan tipe 2 terjadi apabila hipotesis null tidak ditolak ketika hipotesis ini salah.

Adakalanya dijumpai masalah yang memerlukan pengujian lebih dari satu hipotesis pada satu waktu yang dilakukan secara simultan. Pengujian tersebut dinamakan pengujian hipotesis berganda. Sebagai contoh, Misal akan diuji beda mean diantara 5 populasi, maka akan dilakukan sebanyak C_2^5 pengujian beda mean antar populasi. Jika tingkat signifikansi untuk

setiap pengujian α , maka dalam pengujian tersebut telah dilakukan kesalahan sebesar $C_2^5 \alpha$. Kesalahan tipe 1 akan meningkat seiring dengan meningkatnya jumlah pengujian.

Bagaimana halnya dengan masalah biologi? Untuk menguji perbedaan ekspresi gen, akan melibatkan banyak sekali pengujian hingga berjumlah ratusan atau ribuan. Hal ini mengakibatkan meningkatnya besar kesalahan tipe 1. Untuk mengatasi hal itu diperlukan suatu cara untuk mengukur kesalahan tipe 1 secara keseluruhan.

Ada beberapa cara untuk mengukur kesalahan tipe 1 secara keseluruhan pada pengujian hipotesis berganda antara lain *Family Wise Error Rate* (FWER), *False Discovery Rate* (FDR), dan *positif False Discovery Rate* (pFDR). Ketiganya memiliki metode pengujian yang berbeda.

Metode pengujian untuk hipotesis berganda adalah Metode Bonferroni yang merupakan salah satu metode untuk FWER, metode Benjamin-Hochberg untuk FDR, dan metode Storey untuk pFDR.

Dalam tugas akhir ini, akan dibahas cara mengukur kesalahan tipe 1 secara keseluruhan pada pengujian hipotesis berganda disertai penjelasan tentang ketiga metode pengujian untuk hipotesis berganda

1.2 PERUMUSAN MASALAH

Perumusan masalah pada tugas akhir ini adalah bagaimana menguji hipotesis berganda

1.3 TUJUAN

Tujuan dari penulisan adalah membahas metode-metode pengujian untuk hipotesis berganda

1.4 PEMBATASAN MASALAH

Pada tugas akhir ini, pembatasan masalahnya adalah :

1. Statistik uji untuk setiap pengujian hipotesis adalah bersifat independen.
2. π_0 yang menyatakan proporsi H_0 benar di antara semua pengujian yang terkait dengan metode Storey, diasumsikan diketahui.

1.5 SISTEMATIKA PENULISAN

Penulisan tugas akhir yang merupakan hasil studi pustaka ini, dibagi menjadi lima bab, yaitu :

Bab I : Pendahuluan

Membahas tentang latar belakang masalah, perumusan masalah, tujuan penulisan, pembatasan masalah, dan sistematika penulisan.

Bab II : Landasan Teori

Membahas tentang dasar-dasar teori yang digunakan dalam penulisan skripsi, yaitu variabel random, ekspektasi, ekspektasi bersyarat, sampel

random, distribusi statistik terurut, hukum De Morgan, teorema Bayes, pengujian hipotesis, pengujian hipotesis tunggal.

Bab III : Membahas pengujian hipotesis berganda

Bab IV : Aplikasi dari metode-metode pengujian untuk hipotesis berganda

Bab V : Kesimpulan dan saran untuk tugas akhir ini.



BAB II

LANDASAN TEORI

Pada bab ini akan dibahas teori-teori yang mendukung metode-metode pengujian untuk hipotesis berganda, antara lain variabel random, ekspektasi, sampel random, distribusi statistik terurut, hokum de Morgan, teorema Bayes, pengujian hipotesis, dan pengujian hipotesis tunggal.

2.1 VARIABEL RANDOM

Misalkan dilakukan percobaan random dengan ruang sampel \mathcal{E} . Sebuah ruang sampel \mathcal{E} mungkin menggambarkan elemen dari \mathcal{E} yang bukan angka. Misalnya dapat dilihat dalam percobaan random pelemparan sebuah mata uang, ruang sampel yang berkaitan dengan percobaan tersebut adalah $\mathcal{E} = \{c \mid c \text{ adalah muka atau } c \text{ adalah belakang}\}$. Misalkan X adalah sebuah fungsi sedemikian sehingga $X(c) = 0$ jika c adalah muka dan $X(c) = 1$ jika c adalah belakang. Fungsi X adalah variabel random.

Definisi 1

Misalkan suatu percobaan random dengan ruang sampel \mathcal{E} . Suatu fungsi X , yang memetakan setiap element $c \in \mathcal{E}$ ke satu dan hanya satu

bilangan real $x \in \mathcal{A}$ sedemikian sehingga $X(c) = x$, disebut variabel random.

Domain dari X adalah \mathcal{C} dan range dari X adalah $\mathcal{A} = \{x | X(c) = x, c \in \mathcal{C}\}$.

$C \subset \mathcal{C}$ terjadi jika dan hanya jika $A \subset \mathcal{A}$ terjadi, dimana $C = \{c | c \in \mathcal{C}, X(c) \in A\}$, sehingga probabilitas kejadian A sama dengan probabilitas bahwa hasil suatu percobaan berada di C atau maka $\Pr\{X \in A\} = P(C)$, $P(C)$ adalah probabilitas kejadian C terjadi dan P adalah fungsi himpunan probabilitas. Ada dua jenis variabel random yaitu variabel random diskrit dan variabel random kontinu.

Definisi 2

Misalkan X adalah variabel random dengan ruang hasil \mathcal{A} . Misalkan \mathcal{A} merupakan suatu himpunan yang berisi titik-titik berhingga atau dapat dikorespondensikan satu-satu dengan bilangan bulat positif. \mathcal{A} dengan sifat seperti ini disebut himpunan diskrit. Misalkan f suatu fungsi sedemikian sehingga :

1. $f(x) > 0, x \in \mathcal{A}$
2. $\sum_{\mathcal{A}} f(x) = 1$
3. $P(A) = \Pr(X \in A) = \sum_A f(x)$

Maka X disebut variabel random diskrit dan $f(x)$ disebut fungsi kepadatan dari X atau (*Probability Density Function*) yang selanjutnya disingkat p.d.f.

Definisi 3

Misalkan X adalah variabel random dengan ruang hasil \mathcal{A} , $\mathcal{A} \subset \mathbb{R}$ dan f adalah suatu fungsi pada bilangan riil. Jika :

1. $f(x) > 0, \forall x \in \mathcal{A}$

2. $\int_{\mathcal{A}} f(x) dx = 1$

3. $A \subset \mathcal{A}$ berlaku $P(A) = \Pr(X \in A) = \int_A f(x) dx$

dipenuhi, maka X disebut variabel random kontinu dan $f(x)$ disebut fungsi kepadatan dari X atau (*Probability Density Function*) yang selanjutnya disingkat p.d.f

Misalkan variabel random X memiliki probabilitas $P(A)$, $A \subset \mathbb{R}$. Ambil suatu bilangan riil x dan anggap himpunan A adalah himpunan tak terbatas dari $-\infty$ sampai x , titik x termasuk dalam himpunan tersebut. Untuk setiap himpunan berlaku $P(A) = \Pr\{X \in A\} = \Pr\{X \leq x\}$. $F(x) = \Pr\{X \leq x\}$, fungsi $F(x)$ disebut fungsi distribusi atau fungsi distribusi kumulatif dari variabel random X . Karena $F(x) = \Pr\{X \leq x\}$, maka dengan menggunakan p.d.f

$f(x)$, fungsi distribusi dari X dapat dinyatakan dengan $F(x) = \sum_{w \leq x} f(w)$ untuk

variable random diskrit dan $F(x) = \int_{-\infty}^x f(w) dw$ untuk variabel random kontinu.

2.2 EKSPEKTASI

Misalkan X suatu variabel random dengan p.d.f $f(x)$, jika X adalah variabel random diskrit dan $\sum_x |x| f(x)$ konvergen maka ekspektasi dari X :

$E(X) = \sum_x x f(x)$. Jika X adalah variabel random kontinu dan

$\int_{-\infty}^{\infty} |x| f(x) dx$ konvergen maka ekspektasi dari X : $E(X) = \int_{-\infty}^{\infty} x f(x) dx$.

2.2.1 Ekspektasi Bersyarat

Jika X dan Y adalah variabel random diskrit, dengan p.d.f bersama $f(x, y)$ dan p.d.f marjinal $f_X(x)$ dan $f_Y(y)$, maka probabilitas bersyarat dari X diberikan $Y = y$, didefinisikan oleh :

$$f_{X|Y}(x|y) = P\{X = x|Y = y\} = \frac{P\{X = x, Y = y\}}{P\{Y = y\}} = \frac{f(x, y)}{f_Y(y)}, \text{ dimana } P\{Y = y\} > 0$$

atau $f_Y(y) > 0$.

Ekspektasi bersyarat dari X diberikan $Y = y$ dari p.d.f bersyarat di atas adalah

$$\begin{aligned} E[X|Y = y] &= \sum_x x P\{X = x|Y = y\} \\ &= \sum_x x f_{X|Y}(x|y) \end{aligned}$$

Akan ditunjukkan bahwa : $E[E[X|Y]] = E[X]$

$$\begin{aligned}
E[E[X|Y]] &= \sum_y E[X|Y=y]P\{Y=y\} \\
&= \sum_y \sum_x xP\{X=x|Y=y\}P\{Y=y\} \\
&= \sum_y \sum_x x \frac{P\{X=x, Y=y\}}{P\{Y=y\}} P\{Y=y\} \\
&= \sum_y \sum_x xP\{X=x, Y=y\} \\
&= \sum_x x \sum_y P\{X=x, Y=y\} \\
&= \sum_x xP\{X=x\} \\
&= E[X]
\end{aligned}$$

Jika X dan Y adalah variabel random kontinu, dengan p.d.f bersama $f(x, y)$ dan pdf marginal $f_x(x)$ dan $f_y(y)$, maka probabilitas bersyarat dari

X diberikan $Y = y$ didefinisikan oleh : $f_{x|y}(x|y) = \frac{f(x, y)}{f_y(y)}$, dimana $f_y(y) > 0$.

Ekspektasi bersyarat dari X diberikan $Y = y$ dari pdf bersyarat di atas adalah

$$E[X|Y=y] = \int_{-\infty}^{\infty} xf_{x|y}(x|y)dx$$

Akan ditunjukkan bahwa : $E[E[X|Y]] = E[X]$

$$\begin{aligned}
E[E[X|Y]] &= \int_{-\infty}^{\infty} E[X|Y=y]f_y(y)dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf_{x|y}(x|y)f_y(y) dxdy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x \frac{f(x, y)}{f_y(y)} f_y(y) dxdy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x, y) dxdy
\end{aligned}$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x, y) dy dx \\
&= \int_{-\infty}^{\infty} x f_X(x) dx \\
&= E[X]
\end{aligned}$$

2.3 SAMPEL RANDOM

Misalkan X_1, X_2, \dots, X_n menotasikan n buah variabel random yang saling bebas dan masing-masing mempunyai p.d.f yang sama yaitu $f(x)$, artinya p.d.f dari X_1, X_2, \dots, X_n masing-masing adalah

$f_1(x_1) = f(x_1), f_2(x_2) = f(x_2), \dots, f_n(x_n) = f(x_n)$, sehingga p.d.f bersama dari X_1, X_2, \dots, X_n adalah $f(x_1, x_2, \dots, x_n) = f(x_1) f(x_2) \dots f(x_n)$, maka X_1, X_2, \dots, X_n

disebut sampel random dari distribusi dengan p.d.f $f(x)$. Artinya observasi-observasi dari sampel random independen dan mempunyai distribusi yang sama, sering dinotasikan dengan i.i.d (*independent and identically distributed*).

2.4 DISTRIBUSI STATISTIK TERURUT

Misalkan X_1, X_2, \dots, X_n adalah sample random dari distribusi tipe kontinu dengan p.d.f $f(x)$ yang bernilai positif pada interval $a < x < b$.

Kemudian, misalkan Y_1 adalah yang terkecil dari X_i , Y_2 adalah yang terkecil berikutnya dari X_i, \dots, Y_n adalah yang terbesar dari X_i . Sehingga diperoleh $Y_1 < Y_2 < \dots < Y_n$. Maka $Y_i ; i = 1, 2, \dots, n$ disebut statistik terurut ke- i dari sample random X_1, X_2, \dots, X_n , dan p.d.f bersama dari Y_1, Y_2, \dots, Y_n adalah :

$$g(y_1, y_2, \dots, y_n) = (n!)f(y_1)f(y_2)\dots f(y_n), \quad a < y_1 < \dots < y_n < b$$

$$= 0 \quad , \text{ yang lainnya}$$

Untuk $a < y_n < b$, p.d.f marginal dari Y_n adalah :

$$g_n(y_n) = \int_a^{y_n} \dots \int_a^{y_4} \int_a^{y_3} \int_a^{y_2} (n!)f(y_1)f(y_2)\dots f(y_n) dy_1 dy_2 dy_3 \dots dy_{n-1}$$

$$= n! \int_a^{y_n} \dots \int_a^{y_4} \int_a^{y_3} \left(\int_a^{y_2} f(y_1) dy_1 \right) f(y_2)\dots f(y_n) dy_2 dy_3 \dots dy_{n-1}$$

$$= n! \int_a^{y_n} \dots \int_a^{y_4} \int_a^{y_3} F(y_2)f(y_2)\dots f(y_n) dy_2 dy_3 \dots dy_{n-1}$$

$$= n! \int_a^{y_n} \dots \int_a^{y_4} \left(\int_a^{y_3} F(y_2)f(y_2) dy_2 \right) \dots f(y_n) dy_3 \dots dy_{n-1}$$

$$= n! \int_a^{y_n} \dots \left(\int_a^{y_4} \frac{[F(y_3)]^2}{2} f(y_3) dy_3 \right) f(y_n) dy_4 \dots dy_{n-1}$$

Catatan:

Perhatikan $\int F(y_2)f(y_2)dy_2$

$$u = F(y_2) \quad \quad \quad dv = f(y_2)dy_2$$

$$du = F'(y_2)dy_2 \quad \quad \quad v = \int f(y_2)dy_2$$

$$du = f(y_2)dy_2 \quad \quad \quad v = F(y_2)$$

$$\int F(y_2)f(y_2)dy_2 = [F(y_2)]^2 - \int F(y_2)f(y_2)dy_2$$

$$2\int F(y_2)f(y_2)dy_2 = [F(y_2)]^2$$

$$\int F(y_2)f(y_2)dy_2 = \frac{[F(y_2)]^2}{2}$$

Dengan cara yang sama, $\int [F(y_2)]^2 f(y_2)dy_2 = \frac{1}{3}[F(y_2)]^3$ dan

Secara umum : $\int_a^x [F(w)]^{\alpha-1} f(w)dw = \frac{[F(x)]^\alpha}{\alpha}, \alpha > 0$

$$g_n(y_n) = n! \int_a^{y_n} \dots \left(\int_a^{y_4} \frac{[F(y_3)]^2}{2} f(y_3)dy_3 \right) f(y_n)dy_4 \dots dy_{n-1}$$

Dan seterusnya sehingga diperoleh

$$g_n y_n = n! \left(\frac{[F(y_n)]^{n-1}}{(n-1)!} f(y_n) \right) = n [F(y_n)]^{n-1} f(y_n), a < y_n < b$$

= 0 , yang lain

2.5 HUKUM DE MORGAN

Misal S adalah ruang sample dari suatu eksperimen dan A_1, A_2, \dots, A_k adalah peristiwa-peristiwa di dalam S sedangkan $A_1^c, A_2^c, \dots, A_m^c$ adalah komplemen dari peristiwa-peristiwa di dalam S sedemikian sehingga A_1, A_2, \dots, A_k tidak saling asing. Akan ditunjukkan, Hukum De Morgan yang pertama, yaitu :

$$(A_1 \cup A_2 \cup \dots \cup A_m)^c = A_1^c \cap A_2^c \cap \dots \cap A_m^c$$

➤ Akan dibuktikan bahwa : Jika $x \in (A_1 \cup A_2 \cup \dots \cup A_m)^c$ maka

$$x \in (A_1^c \cap A_2^c \cap \dots \cap A_m^c):$$

Misalkan $x \in (A_1 \cup A_2 \cup \dots \cup A_m)^c$ berarti $x \notin (A_1 \cup A_2 \cup \dots \cup A_m)$, Hal ini

berarti bahwa x tidak terdapat pada salah satu $A_i; i = 1, 2, \dots, m$. Oleh

karena itu, $x \in A_i^c; i = 1, 2, \dots, m$ sehingga $x \in (A_1^c \cap A_2^c \cap \dots \cap A_m^c)$

➤ Akan dibuktikan bahwa : Jika $x \in (A_1^c \cap A_2^c \cap \dots \cap A_m^c)$ maka

$$x \in (A_1 \cup A_2 \cup \dots \cup A_m)^c:$$

Misalkan $x \in (A_1^c \cap A_2^c \cap \dots \cap A_m^c)$ berarti $x \in A_i^c; i = 1, 2, \dots, m$ dan

$x \notin A_i; i = 1, 2, \dots, m$. Oleh karena itu, $x \notin (A_1 \cup A_2 \cup \dots \cup A_m)$ sehingga

$$x \in (A_1 \cup A_2 \cup \dots \cup A_m)^c.$$

Karena $x \in (A_1 \cup A_2 \cup \dots \cup A_m)^c$ jika dan hanya jika $x \in (A_1^c \cap A_2^c \cap \dots \cap A_m^c)$,

maka $(A_1 \cup A_2 \cup \dots \cup A_m)^c = A_1^c \cap A_2^c \cap \dots \cap A_m^c$.

Selanjutnya, akan ditunjukkan Hukum De Morgan yang kedua :

$$(A_1 \cap A_2 \cap \dots \cap A_m)^c = A_1^c \cup A_2^c \cup \dots \cup A_m^c$$

Bukti:

Untuk membuktikan hukum De Morgan kedua, gunakan hukum De

Morgan yang pertama untuk mendapatkan :

$$(A_1^c \cup A_2^c \cup \dots \cup A_m^c)^c = (A_1^c)^c \cap (A_2^c)^c \cap \dots \cap (A_m^c)^c,$$

karena $(A^c)^c = A$, maka $(A_1^c)^c \cap (A_2^c)^c \cap \dots \cap (A_m^c)^c = A_1 \cap A_2 \cap \dots \cap A_m$.

Jadi, $A_1 \cap A_2 \cap \dots \cap A_m = (A_1^c \cup A_2^c \cup \dots \cup A_m^c)^c$

Dan $(A_1 \cap A_2 \cap \dots \cap A_m)^c = A_1^c \cup A_2^c \cup \dots \cup A_m^c$

2.6 TEOREMA BAYES

2.6.1 Probabilitas dan Partisi

Misal S adalah ruang sample dari suatu eksperimen dan A_1, A_2, \dots, A_k adalah peristiwa-peristiwa di dalam S yang membentuk partisi di dalam S sedemikian sehingga A_1, A_2, \dots, A_k saling asing dan $\bigcup_{i=1}^k A_i = S$. Jika k peristiwa A_1, A_2, \dots, A_k membentuk partisi di dalam S dan misalkan B adalah sebarang peristiwa di dalam S , maka $A_1 \cap B, A_2 \cap B, \dots, A_k \cap B$ membentuk partisi di dalam B . Sehingga dapat kita tulis $B = (A_1 \cap B) \cup (A_2 \cap B) \cup \dots \cup (A_k \cap B)$

Selanjutnya karena peristiwa-peristiwa di ruas kanan saling asing, maka

$P(B) = \sum_{i=1}^k P(A_i \cap B)$. Jika $P(A_i) > 0$ untuk $i=1, 2, \dots, k$ maka

$P(A_i \cap B) = P(A_i)P(B|A_i)$. Sehingga didapat $P(B) = \sum_{i=1}^k P(A_i)P(B|A_i)$.

Sekarang akan diberikan teorema Bayes

2.6.2 Teorema Bayes

Misal peristiwa-peristiwa A_1, A_2, \dots, A_k membentuk partisi di dalam ruang sample sedemikian sehingga $P(A_i) > 0$; $i = 1, 2, \dots, k$ dan misalkan B sebarang peristiwa sedemikian sehingga $P(B) > 0$. Maka untuk $i=1, 2, \dots, k$,

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^k P(A_j)P(B|A_j)}$$

Teorema Bayes memberikan aturan sederhana untuk menghitung probabilitas bersyarat peristiwa A_i diberikan B terjadi, jika masing-masing probabilitas tak bersyarat A_i dan probabilitas bersyarat B diberikan A_i terjadi diketahui.

2.7 PENGUJIAN HIPOTESIS

Selain penaksiran parameter, pengujian hipotesis merupakan salah satu bagian yang penting dalam statistika inferensi. Statistika inferensi adalah metode-metode analisis terhadap sampel untuk mendapatkan pendugaan atau penarikan kesimpulan tentang keseluruhan informasi dalam populasi. Pengujian dilakukan terhadap parameter populasi dalam bentuk hipotesis statistik.

Berikut ini adalah langkah-langkah dari pengujian :

1. Membuat pernyataan hipotesis.

2. Statistik uji disesuaikan berdasarkan informasi dari distribusi dan hipotesis dalam pengujian.
3. Aturan keputusan dibuat berdasarkan sampel yang diambil dari populasi dengan menghitung nilai statistik uji untuk memutuskan apakah menerima atau menolak hipotesis.

Untuk menjelaskan langkah-langkah di atas, berikut ini diterangkan beberapa konsep dalam pengujian hipotesis.

2.7.1 Hipotesis Statistik

Definisi 2.7.1

Hipotesis statistik adalah pernyataan atau dugaan mengenai satu atau lebih parameter

Definisi 2.7.2

Pengujian hipotesis adalah suatu aturan yang mengacu pada keputusan menerima atau menolak hipotesis statistik yang diuji, berdasarkan nilai yang diperoleh dari sampel.

Untuk membentuk sebuah pengujian statistika, biasanya beberapa informasi pendukung telah diketahui terlebih dahulu. Berdasarkan informasi pendukung tersebut, pernyataan yang telah diketahui dan pernyataan yang

akan diuji dirangkum dalam hipotesis statistik, yaitu hipotesis null dan hipotesis alternatif.

Hipotesis null, biasanya disimbolkan dengan H_0 , menyatakan pernyataan yang telah diketahui sebelumnya. Sedangkan hipotesis alternatif, biasa disimbolkan dengan H_1 , menyatakan pernyataan yang akan dicapai dalam pengujian statistika.

Sebagai contoh pengaruh penggunaan obat jenis baru. Hipotesis null adalah rata-rata lama waktu sembuh mengkonsumsi obat jenis baru sama dengan rata-rata lama waktu sembuh mengkonsumsi obat jenis lama. Maka dapat ditulis H_0 : tidak terdapat perbedaan rata-rata lama waktu sembuh antara kedua obat. Sedangkan hipotesis alternatif adalah rata-rata lama waktu sembuh mengkonsumsi obat jenis baru lebih kecil daripada rata-rata lama waktu sembuh mengkonsumsi obat jenis lama. Maka dapat ditulis H_1 : terdapat perbedaan rata-rata lama waktu sembuh antara kedua obat.

Pengujian hipotesis dilakukan bila ada dugaan yang berbeda dari pernyataan yang telah diketahui sebelumnya. Oleh karena itu, pengujian hipotesis selalu dilakukan terhadap hipotesis null, sehingga kesimpulan akhir yang diperoleh adalah menolak H_0 atau tidak menolak H_0 , bukan menerima H_1 atau menolak H_1 . Jika disimpulkan bahwa H_0 tidak ditolak tidak berarti H_0 adalah benar, hanya dapat disimpulkan bahwa berdasarkan informasi dari sampel tidak terdapat cukup bukti untuk menolak H_0 . Penolakan H_0 menyarankan bahwa hipotesis alternatif mungkin benar.

Apabila hipotesis menyatakan distribusi populasi secara jelas, maka hipotesis disebut hipotesis sederhana. Contoh hipotesis sederhana :

$$H_0 : \theta = 75$$

$$H_1 : \theta > 75$$

Sedangkan jika distribusi populasi tidak dinyatakan secara jelas, maka hipotesis tersebut disebut hipotesis komposit. Contoh hipotesis komposit :

$$H_0 : \theta \leq 75$$

$$H_1 : \theta > 75$$

Dalam pengujian hipotesis, terdapat dua jenis kesalahan yang mungkin terjadi. Kesalahan tipe 1 terjadi apabila hipotesis null ditolak ketika hipotesis ini benar. Kesalahan tipe 2 terjadi apabila hipotesis null tidak ditolak ketika hipotesis ini salah.

Sebagai contoh pengaruh penggunaan obat jenis baru terhadap rata-rata lama waktu sembuh dibandingkan menggunakan obat yang telah ada. Hipotesis null, H_0 : tidak terdapat perbedaan rata-rata lama waktu sembuh antara obat baru dengan obat yang telah ada. Hipotesis alternatif, H_1 : terdapat perbedaan rata-rata lama waktu sembuh antara obat baru dengan obat yang telah ada. Kesalahan tipe 1 terjadi apabila disimpulkan bahwa rata-rata lama waktu sembuh obat jenis baru berbeda dari yang telah ada, padahal kenyataannya tidak demikian. Kesalahan tipe 2 muncul jika disimpulkan bahwa rata-rata lama waktu sembuh obat jenis baru tidak berbeda dari yang telah ada, padahal kenyataannya berbeda.

2.7.2 Statistik Uji

Statistik uji adalah statistik yang digunakan untuk membantu membuat kesimpulan menolak atau menerima hipotesis (pernyataan) pada pengujian hipotesis.

Pemilihan statistik uji bergantung kepada informasi distribusi dan hipotesis dalam pengujian. Berikut adalah definisi-definisi yang dikenal dalam pengujian hipotesis.

Definisi 2.7.3

Suatu ruang sampel dibagi menjadi dua daerah, yaitu daerah penolakan H_0 dan daerah tidak menolak H_0 . Selanjutnya daerah penolakan H_0 disebut daerah kritis.

Sesuai dengan aturan keputusan, jika nilai statistik uji jatuh pada daerah kritis, maka membawa keputusan pada penolakan H_0 . Sebaliknya, bila nilai statistik uji tidak jatuh pada daerah kritis atau berada pada daerah tidak menolak H_0 , maka keputusannya adalah tidak menolak H_0 .

Definisi 2.7.4

Nilai kritis untuk hipotesis adalah suatu nilai batas antara daerah kritis dan daerah tidak menolak H_0 . Nilai kritis untuk setiap pengujian hipotesis bergantung kepada tingkat signifikansi pengujian hipotesis yang dipilih.

Definisi 2.7.5

Power function dari suatu pengujian adalah suatu fungsi dari parameter yang menghasilkan probabilitas bahwa titik sampel jatuh dalam daerah kritis atau suatu fungsi yang menghasilkan probabilitas penolakan H_0 . Nilai dari *power function* pada sebuah titik parameter disebut *power* pengujian dari titik tersebut.

Power dari pengujian hipotesis adalah probabilitas nilai sampel berada pada daerah kritis, yaitu probabilitas penolakan H_0 .

Definisi 2.7.6

Tingkat signifikansi dari suatu pengujian hipotesis adalah probabilitas menolak H_0 ketika H_0 benar atau nilai maksimum *power function* dari suatu pengujian ketika H_0 benar

Definisi 2.7.7

Nilai probabilitas (*p-value*) dari suatu pengujian hipotesis adalah probabilitas yang diperoleh dari nilai statistik uji, dihitung ketika H_0 benar.

Untuk memutuskan menolak atau tidak menolak H_0 , nilai *p-value* dibandingkan terhadap nilai signifikansi. Jika nilai *p-value* kurang dari tingkat signifikansi, hipotesis null akan ditolak.

2.7.3 Aturan Keputusan

Misalkan hipotesis statistik yang akan diuji telah didefinisikan, selanjutnya diperlukan suatu aturan untuk mengambil tindakan menolak atau menerima hipotesis tersebut. Keputusan menolak hipotesis null dilakukan bila statistik uji berada di daerah kritis, sebaliknya bila statistik uji berada di daerah tidak menolak H_0 maka tidak cukup alasan untuk menolak H_0 .

Bila statistik uji jatuh pada daerah kritis, sebenarnya yang terjadi adalah nilai *p-value* lebih kecil dari tingkat signifikansi yang dipilih. Sehingga cukuplah dilakukan tindakan menolak atau tidak menolak H_0 hanya berdasarkan membandingkan nilai *p-value* dengan tingkat signifikansi yang dipilih. Bila *p-value* lebih kecil daripada tingkat signifikansi, keputusan adalah menolak H_0 .

2.8 PENGUJIAN HIPOTESIS TUNGGAL

Misalkan ingin diuji sebuah hipotesis null H_0 versus hipotesis alternatif H_1 , Hipotesisnya adalah :

$$H_0 : \theta = \theta_0$$

$$H_1 : \theta \neq \theta_0$$

Dimana : θ adalah suatu parameter

θ_0 adalah suatu nilai dari parameter yang dihipotesiskan.

Misalkan dilakukan pengujian dengan statistik uji T dan diberikan sebuah daerah kritis Γ , maka H_0 ditolak ketika $T \in \Gamma$ dan H_0 tidak ditolak ketika $T \notin \Gamma$. Kesalahan tipe 1 terjadi ketika $T \in \Gamma$ padahal H_0 benar dan kesalahan tipe 2 terjadi ketika $T \notin \Gamma$ padahal H_0 salah.

Berikut disajikan tabel probabilitas kejadian-kejadian dari suatu pengujian hipotesis tunggal:

Tabel 2.1 : Probabilitas kejadian-kejadian dari pengujian tunggal

	H_0 tidak ditolak	H_0 ditolak
H_0 benar	$1-\alpha$	α
H_0 salah	β	$1-\beta$

Dimana :

α adalah probabilitas menolak H_0 ketika H_0 benar (probabilitas melakukan kesalahan tipe1)

β adalah probabilitas tidak menolak H_0 ketika H_0 salah (probabilitas melakukan kesalahan tipe2)

$1-\alpha$ adalah probabilitas tidak menolak H_0 ketika H_0 benar (keputusan benar)

$1-\beta$ adalah probabilitas menolak H_0 ketika H_0 salah (keputusan benar)

Pemilihan daerah kritis terbaik, adalah memilih daerah kritis sedemikian sehingga probabilitas kesalahan tipe1 $\leq \alpha$, kemudian diantara daerah kritis-daerah kritis ini, dipilihlah satu yang meminimumkan probabilitas

kesalahan tipe2 (β). Artinya, memilih daerah kritis yang memaksimumkan power ($1-\beta$) dengan mempertahankan probabilitas kesalahan tipe1 pada suatu level tertentu.



BAB III

METODE-METODE PENGUJIAN UNTUK HIPOTESIS BERGANDA

3.1 PENGERTIAN PENGUJIAN HIPOTESIS BERGANDA

Pengujian hipotesis berganda adalah pengujian lebih dari satu hipotesis pada satu waktu yang dilakukan secara simultan. Misalkan dilakukan m pengujian hipotesis identik, hipotesisnya adalah :

$$H_{0i} : \theta = \theta_0$$

$$H_{1i} : \theta \neq \theta_0$$

Dimana : H_{0i} adalah H_0 untuk hipotesa ke i ; $i=1,2,\dots,m$

H_{1i} adalah H_1 untuk hipotesa ke i ; $i=1,2,\dots,m$

θ adalah suatu parameter

θ_0 adalah suatu nilai dari parameter yang dihipotesiskan

Kumpulan hipotesis-hipotesis di atas, dinamakan *family hipotesis*.

Contohnya, seorang guru Sosiologi yang mengajar di tiga SMA di Jakarta menerapkan dua metode pengajaran yang berbeda di setiap SMA yaitu metode pengajaran pasif dan metode pengajaran aktif. Ingin dilihat apakah ada perbedaan rata-rata nilai ulangan umum mata pelajaran Sosiologi diantara dua metode pengajaran untuk masing-masing SMA.

Hipotesisnya adalah $H_{01} : \mu_{11} = \mu_{12}$, $H_{02} : \mu_{21} = \mu_{22}$, dan $H_{03} : \mu_{31} = \mu_{32}$
 $H_{11} : \mu_{11} \neq \mu_{12}$, $H_{12} : \mu_{21} \neq \mu_{22}$, dan $H_{13} : \mu_{31} \neq \mu_{32}$

Dimana :

μ_{11} adalah rata-rata nilai ulangan umum SMA 1 metode pengajaran pasif.

μ_{12} adalah rata-rata nilai ulangan umum SMA 1 metode pengajaran aktif.

:

μ_{32} adalah rata-rata nilai ulangan umum SMA 3 metode pengajaran aktif.

Pada pengujian hipotesis berganda, karena pengujian-pengujian dilakukan pada satu waktu secara simultan, maka akan terjadi penggandaan nilai probabilitas kesalahan tipe 1 (menolak H_0 ketika H_0 benar). Untuk mengatasi hal itu diperlukan suatu cara untuk mengukur kesalahan tipe 1 secara keseluruhan. Ada beberapa cara untuk mengukur kesalahan tipe 1 secara keseluruhan pada pengujian hipotesis berganda antara lain *Family Wise Error Rate* (FWER), *False Discovery Rate* (FDR), dan *positif False Discovery Rate* (pFDR). Ketiganya memiliki metode pengujian yang berbeda.

Metode pengujian untuk hipotesis berganda adalah Metode Bonferroni yang merupakan salah satu metode untuk FWER, metode Benjamin-Hochberg untuk FDR, dan metode Storey untuk pFDR.

3.2. FAMILY WISE ERROR RATE (FWER)

Probabilitas melakukan kesalahan tipe 1 secara keseluruhan diukur dengan mempertimbangkan keberadaan kesalahan tipe 1 dalam suatu family hipotesis yaitu probabilitas melakukan paling sedikit satu kesalahan tipe 1 diantara seluruh pengujian dalam *family hipotesis*. Hal inilah yang disebut *Family Wise Error Rate* (FWER). Apabila pengujian hipotesis berganda dilakukan dengan tingkat signifikansi α , maka probabilitas kesalahan tipe 1 secara keseluruhan dalam *family hipotesis* perlu dikontrol sedemikian sehingga $FWER \leq \alpha$.

Misalkan pada pengujian hipotesis berganda dilakukan m pengujian secara independen, masing-masing dengan tingkat signifikansi α dan misalkan A_i adalah kejadian bahwa H_0 ke i tidak ditolak, A_i' adalah kejadian bahwa H_0 ke i ditolak, dimana $i=1,2,\dots,m$. Jika setiap H_0 benar, $P(A_i) = 1-\alpha$ dan $P(A_i') = \alpha$, maka FWER atau probabilitas melakukan paling sedikit satu kesalahan tipe 1 di antara m pengujian adalah :

$$P(A_1' \cup A_2' \cup \dots \cup A_m') = 1 - P\left[(A_1' \cup A_2' \cup \dots \cup A_m')'\right]$$

Dari hukum de Morgan, diketahui bahwa

$$(A_1' \cup A_2' \cup \dots \cup A_m')' = A_1 \cap A_2 \cap \dots \cap A_m$$

$$\begin{aligned} 1 - P\left[(A_1' \cup A_2' \cup \dots \cup A_m')'\right] &= 1 - P(A_1 \cap A_2 \cap \dots \cap A_m) \\ &= 1 - [P(A_1) \cdot P(A_2) \dots P(A_m)] \\ &= 1 - [(1-\alpha) \cdot (1-\alpha) \dots (1-\alpha)] \\ &= 1 - (1-\alpha)^m \end{aligned}$$

Berikut ini tabel nilai FWER pada pengujian hipotesis berganda dengan m pengujian. Andai dipilih tingkat signifikan $\alpha=0.05$.

Tabel 3.1 : Nilai FWER pada pengujian hipotesis berganda dengan m pengujian

m	$1 - (1 - \alpha)^m$
1	0.05
2	0.098
3	0.143
4	0.185
5	0.226
10	0.401
20	0.642

Hal ini menunjukkan bahwa pada pengujian hipotesis berganda terjadi penggandaan nilai probabilitas kesalahan tipe 1, sehingga kenyataannya nilai FWER lebih besar dari α dan FWER akan meningkat seiring dengan meningkatnya jumlah pengujian.

3.2.1 Metode Bonferroni

Untuk mengontrol FWER, diperlukan suatu metode sedemikian sehingga $\text{FWER} \leq \alpha$, yaitu salah satunya adalah metode Bonferroni.

Berikut ini akan dijelaskan penggunaan dari metode Bonferroni.

Teorema : Misalkan A_1, A_2, \dots, A_m mewakili m kejadian , A_i adalah kejadian bahwa H_0 ke i tidak ditolak, A_i' adalah kejadian bahwa H_0 ke i ditolak, dimana $i=1,2,3,4$. Jika setiap H_0 benar, $P(A_i) = 1-\alpha$ dan $P(A_i') = \alpha$, maka

Pertidaksamaan Bonferroni :

$$P(A_1 \cap A_2 \cap \dots \cap A_m) \geq 1 - \sum_{i=1}^m P(A_i')$$

Bukti : Dari hukum de Morgan , diketahui bahwa

$$(A_1 \cap A_2 \cap \dots \cap A_m)' = A_1' \cup A_2' \cup \dots \cup A_m'$$

Dari sifat probabilitas, $P(A_1' \cup A_2' \cup \dots \cup A_m') \leq \sum_{i=1}^m P(A_i')$, maka

$$\begin{aligned} P(A_1 \cap A_2 \cap \dots \cap A_m) &= 1 - P[(A_1 \cap A_2 \cap \dots \cap A_m)'] \\ &= 1 - P(A_1' \cup A_2' \cup \dots \cup A_m') \\ &\geq 1 - \sum_{i=1}^m P(A_i') \end{aligned}$$

Jadi, $P(A_1 \cap A_2 \cap \dots \cap A_m) \geq 1 - \sum_{i=1}^m \alpha = 1 - m\alpha$

Sehingga probabilitas melakukan paling sedikit satu kesalahan tipe 1 di antara m pengujian adalah :

$$\begin{aligned} P(A_1' \cup A_2' \cup \dots \cup A_m') &= 1 - [P(A_1 \cap A_2 \cap \dots \cap A_m)] \\ &= 1 - [P(A_1 \cap A_2 \cap \dots \cap A_m)] \\ &\leq 1 - \left[1 - \sum_{i=1}^m \alpha\right] = \sum_{i=1}^m \alpha = m\alpha \end{aligned}$$

Jadi, dengan pengujian hipotesis berganda, diperoleh kesalahan tipe 1 secara keseluruhan mendekati $m\alpha$.

Berdasarkan pertidaksamaan Bonferroni, untuk mengontrol kesalahan tipe 1 secara keseluruhan pada tingkat signifikansi α , pada masing-masing pengujian dipilih tingkat signifikansi $= \alpha/m$, dimana m adalah banyaknya

pengujian, sehingga
$$P(A_1' \cup A_2' \cup \dots \cup A_m') \leq 1 - \left[1 - \sum_{i=1}^m \alpha/m \right] = \sum_{i=1}^m \alpha/m = \alpha.$$

Jadi, setiap pengujian dipilih tingkat signifikansinya sama dengan α/m , dengan harapan $FWER \leq \alpha$.

Pemilihan tingkat signifikansi untuk setiap pengujian senilai α/m , akan menyebabkan tingkat signifikansi untuk setiap pengujian semakin kecil. Sehingga power untuk setiap pengujian juga akan semakin kecil. Meningkatnya jumlah pengujian berakibatnya mengecilnya power pengujian. Hal inilah, yang menjadi kelemahan dari metode Bonferroni.

3.3 **FALSE DISCOVERY RATE (FDR)**

Untuk mengatasi kelemahan dari FWER, berdasarkan penjelasan di sub bab 3.2, khususnya ketika jumlah pengujian meningkat, diperlukan suatu ukuran kesalahan tipe 1 yang lebih baik. Kali ini, jumlah kesalahan tipe 1 akan diperkirakan dahulu, dengan perkiraan ini diharapkan pengujian akan mempunyai power yang lebih baik dibandingkan metode Bonferroni.

Untuk lebih jelasnya diberikan uraian sbb:

Misalkan terdapat m pengujian hipotesis, m_0 adalah banyaknya H_0 yang benar dan R adalah banyaknya hipotesis yang ditolak. Kejadian-kejadian yang mungkin akan terjadi, disajikan dalam tabel berikut :

Tabel 3.2 : Kejadian-kejadian yang mungkin akan terjadi dalam m pengujian

	H_0 tidak ditolak	H_0 ditolak	Total
H_0 benar	U	V	m_0
H_0 salah	T	S	m_1
	W	R	m

Dimana :

R = Banyaknya H_0 yang ditolak

W = Banyaknya H_0 yang tidak ditolak

m_0 = Banyaknya H_0 benar

m_1 = Banyaknya H_0 salah

V = Banyaknya kesalahan jenis 1 terjadi (menolak H_0 ketika H_0 benar)

S = Banyaknya H_0 ditolak ketika H_0 salah

T = Banyaknya kesalahan tipe 2 terjadi (tidak menolak H_0 ketika H_0 salah)

U = Banyaknya H_0 tidak ditolak ketika H_0 benar

R, W adalah variabel random terobservasi

U, V, S, T adalah variabel random tidak terobservasi

Proporsi banyaknya kesalahan tipe1 terjadi (menolak H_0 ketika H_0 benar) di antara hipotesis-hipotesis yang ditolak dapat ditunjukkan oleh variable random $Q=V/(V+S)$. Q adalah variabel random tidak terobservasi, sama halnya dengan tidak diketahuinya nilai dari V dan S . Kemudian probabilitas melakukan kesalahan tipe 1 secara keseluruhan dalam *family hipotesis* disebut *False Discovery Rate (FDR)* dan didefinisikan sebagai:

$$E[Q] = E\left[\frac{V}{(V+S)}\right] = E\left[\frac{V}{R}\right]$$

Pada definisi diatas dapat terjadi kondisi :

Tidak ada H_0 yang ditolak, hal ini artinya $R=0$ sehingga $V=0$, agar Q terdefinisi maka perlu didefinisikan $Q=0$. Sebaliknya bila ada H_0 ditolak , maka $R>0$ sehingga $Q=V/(V+S)=V/R$.

Jadi dari definisi diatas maka nilai harapan dari proporsi kesalahan penolakan di antara hipotesis yang ditolak didefinisikan sebagai berikut :

$$\text{FDR} = E\left[\frac{V}{R}\right], R > 0$$

$$0, R = 0$$

Karena ketika $R=0$ menyebabkan $E\left[\frac{V}{R}\right] = 0$, maka FDR dapat pula

didefinisikan sebagai:

$$E\left[\frac{V}{R \vee 1}\right] \text{ dimana } R \vee 1 = \max(R, 1)$$

Pengaruh dari $R \vee 1$ adalah menyatakan bahwa $E\left[\frac{V}{R}\right] = 0$ ketika $R=0$ dan $V=0$. Dalam suatu pengujian, ada kemungkinan terjadinya $R=0$ walaupun diberikan $\max(R,1)$. Sehingga perlu menaksir terjadinya $R>0$ atau $\Pr(R > 0)$.

Sehingga definisi FDR dapat pula dinyatakan sebagai :

$$E\left[\frac{V}{R \vee 1}\right] = E\left[\frac{V}{R} \mid R > 0\right] \cdot \Pr(R > 0), \text{ dimana } R \vee 1 = \max(R, 1).$$

3.3.1 Perbandingan Antara FWER dan FDR

Berikut akan dibandingkan antara FWER dan FDR dalam berbagai kondisi :

Kasus 1 : Jika semua H_0 benar ($m = m_0$),

ketika $m=m_0$, berarti $S=0, V=R$

jika $V > 0$, maka $\frac{V}{R} = 1$, jadi $E\left[\frac{V}{R} \mid R > 0\right] = 1$, sehingga

$$\text{FDR} = 1 \cdot \Pr(R > 0) = \Pr(V \geq 1) = \text{FWER}$$

Jadi, $\text{FDR} = \text{FWER}$

Kasus 2 : Jika $m_0 < m$, maka

$$\left. \begin{array}{l} \text{jika } V=0, S>0, \text{ maka } \frac{V}{R} = 0 \\ \text{jika } V>0, S=0, \text{ maka } \frac{V}{R} = 1 \\ \text{jika } V>0, S>0, \text{ maka } \frac{V}{R} < 1 \end{array} \right\} \text{ jadi, } E\left(\frac{V}{R} \mid R > 0\right) \leq 1$$

$$\begin{aligned} \text{FDR} &= E\left(\frac{V}{R} \mid R > 0\right) \cdot \Pr(R > 0) \\ &\leq 1 \cdot \Pr(R > 0) , \Pr(R > 0) = \Pr(V \geq 1) \\ \text{FDR} &\leq \Pr(V \geq 1) \end{aligned}$$

jadi, $\text{FDR} \leq \text{FWER}$

Dari pembuktian diatas ditunjukkan:

Ketika $m_0 = m$ yang artinya tak ada H_0 yang salah maka kesalahan tipe1 keseluruhan dalam family hipotesis pada FDR sama dengan dalam FWER. Ketika $m_0 < m$ yang artinya ada H_0 yang salah maka kesalahan tipe1 keseluruhan dalam family hipotesis pada FDR akan lebih kecil daripada FWER. Maka disimpulkan bahwa FDR lebih baik daripada FWER. Dengan perkataan lain dapat disimpulkan bahwa , prosedur yang mengontrol FWER juga mengontrol FDR dan prosedur yang hanya mengontrol FDR mempunyai power yang lebih baik daripada prosedur yang mengontrol FWER. Untuk lebih jelasnya, pada sub-bab selanjutnya akan dibahas mengenai prosedur yang mengontrol FDR.

3.3.2 Metode Benjamin dan Hochberg

Benjamin dan Hochberg (1995) memperkenalkan suatu metoda atau prosedur yang bertujuan memperbaiki metoda Bonferroni. Prosedur Benjamin dan Hochberg mengontrol FDR pada tingkat signifikansi α :

Misalkan pada pengujian hipotesis berganda dilakukan m pengujian dengan hipotesis masing-masing H_1, H_2, \dots, H_m dan p -value masing-masing P_1, P_2, \dots, P_m . Selanjutnya p -value tersebut diurutkan dari yang terkecil ke terbesar yakni $P_{(1)} \leq P_{(2)} \leq \dots \leq P_{(m)}$ dimana $P_{(i)}$ adalah p -value yang bersesuaian dengan hipotesis $H_{(i)}$. Jika k adalah bilangan i terbesar yang memenuhi $P_{(i)} \leq \frac{i}{m} \alpha$, maka Hipotesis Null (H_0) pada hipotesis $H_{(1)}, H_{(2)}, \dots, H_{(k)}$ ditolak. (1)

Untuk membuktikan prosedur (1) mengontrol FDR pada tingkat signifikansi α , diberikan teorema berikut:

Teorema :

Untuk statistik uji independent dan sebarang konfigurasi dari H_0 salah, prosedur (1) mengontrol FDR pada q^*

Pembuktian teorema dibantu dengan adanya lemma berikut :

Lemma :

Untuk sebarang $0 \leq m_0 \leq m$ yang menyatakan banyaknya *independent* p -value untuk hipotesis H_0 yang benar, dan $m_1 = m - m_0$ adalah banyaknya p -value untuk hipotesis H_0 yang salah, prosedur (1) memenuhi pertidaksamaan :

$$E \left[Q \mid P_{m_0+1} = p_1, P_{m_0+2} = p_2, \dots, P_{m_0+m_1} = p_{m_1} \right] \leq \frac{m_0}{m} q^* \quad (2)$$

dimana $E[Q] = E\left[\frac{V}{R} | R > 0\right] \Pr(R > 0)$.

$P_{m_0+1}, P_{m_0+2}, \dots, P_m$ menyatakan p -value untuk hipotesis H_0 yang salah.

Pertidaksamaan tersebut ada dengan mengabaikan distribusi dari p -value untuk hipotesis H_0 yang salah.

Dengan membuktikan lemma, didapatkan $E[Q] \leq \frac{m_0}{m} q^* \leq q^*$ dan FDR dapat dikontrol.

Pembuktian Lemma :

Pembuktian lemma dilakukan dengan cara induksi : $\forall m$ berlaku ... (2)

1) Untuk $m=1$, terdapat 2 kondisi yakni ketika $m_0=0$ dan $m_0=1$:

Jika $m_0=0$ maka $Q=0$ dan $E[Q | P_1 = p_1] = 0 \leq \frac{m_0}{m} q^*$, pertidaksamaan (2) terpenuhi.

Jika $m_0=1$ maka $E[Q | P_2 = p_1] \leq \frac{m_0}{m} q^*$, pertidaksamaan (2) terpenuhi.

Jadi, untuk $m=1$ lemma (2) benar.

2) Misalkan untuk m berlaku :

$$E\left[Q \mid P_{m_0+1} = p_1, P_{m_0+2} = p_2, \dots, P_{m_0+m_1} = p_{m_1}\right] \leq \frac{m_0}{m} q^* \text{ adalah benar,}$$

3) Akan ditunjukkan bahwa untuk $m+1$ lemma juga benar, yaitu :

$$E\left[Q \mid P_{m_0+1} = p_1, P_{m_0+2} = p_2, \dots, P_{m_0+m_1} = p_{m_1}\right] \leq \frac{m_0}{m+1} q^*.$$

Ketika $m_0=0$, semua H_0 salah, maka $Q = 0$ dan

$$E\left[Q \mid P_1 = p_1, P_2 = p_2, \dots, P_{m_1=m} = p_m\right] = 0 \leq \frac{m_0}{m+1} q^* , \text{ benar.}$$

Ketika $m_0>0$, tidak semua H_0 salah:

Misalkan $P'_1, P'_2, \dots, P'_{m_0}$ adalah variabel random yang menyatakan p -value

untuk H_0 benar . Selanjutnya urutkan p -value $P'_{(1)} \leq P'_{(2)} \leq \dots \leq P'_{(m_0)}$ dan

$P'_{(m_0)}$ adalah p -value yang paling besar. $P'_{(i)}$; $i = 1, 2, \dots, m_0$ adalah variabel

random independen berdistribusi Uniform (0,1) dan $P'_{(1)}, P'_{(2)}, \dots, P'_{(m_0)}$ adalah p -

value untuk hipotesis $H_{(1)}, H_{(2)}, \dots, H_{(m_0)}$. Misalkan $P_{m_0+1} = p_1, \dots, P_m = p_{m_1}$

menyatakan p -value untuk H_0 salah, diurutkan dari yang terkecil ke

terbesar $p_1 \leq p_2 \leq \dots \leq p_{m_1}$. $P_{(m_0+1)}, \dots, P_{(m)}$ adalah p -value untuk hipotesis

$H_{(m_0+1)}, \dots, H_{(m)}$.

Maka, dengan bersyarat $P'_{(m_0)} = p$, diperoleh

$$\begin{aligned} E\left[Q \mid P_{m_0+1} = p_1, \dots, P_m = p_{m_1}\right] &= E\left[E\left[\left(Q \mid P_{m_0+1} = p_1, \dots, P_m = p_{m_1}\right) \mid P'_{(m_0)} = p\right]\right] \\ &= \int_0^1 E\left[Q \mid P'_{(m_0)} = p, P_{m_0+1} = p_1, \dots, P_m = p_{m_1}\right] f_{P'_{(m_0)}}(p) dp \end{aligned}$$

Akan dicari pdf marginal dari $P'_{(m_0)}$ yaitu $f_{P'_{(m_0)}}(p)$:

pdf dari $P'_{(m_0)}$, dimana $i=1,2,\dots,m_0$ adalah:

$$f(p) = \frac{1}{1-0} = 1, \quad 0 \leq p \leq 1$$

$$= 0, \quad \text{yang lain}$$

Fungsi distribusi dari $P'_{(m_0)}$ adalah

$$F(p) = 0, \quad p < 0$$

$$= \frac{p-0}{1-0} = p, \quad 0 < p < 1$$

$$= 1, \quad p > 1$$

Jadi, pdf marginal dari $P'_{(m_0)}$ adalah

$$f_{P'_{m_0}}(p) = m_0 [F(p)]^{m_0-1} f(p), \quad 0 < p < 1$$

$$= m_0 p^{m_0-1} \cdot 1 = m_0 p^{m_0-1}$$

$$= 0, \quad \text{yang lainnya}$$

Kemudian, akan ditunjukkan:

$$E\left[Q \mid P_{m_0+1} = p_1, P_{m_0+2} = p_2, \dots, P_{m_0+m_1} = p_{m_1}\right] \leq \frac{m_0}{m+1} q^* \text{ benar}$$

jika j_0 adalah bilangan j terbesar, dimana $0 \leq j \leq m_1$, yang memenuhi :

$$p_j \leq \frac{m_0 + j}{m+1} q^* \tag{4}$$

maka Hipotesis null $H_{(1)}, \dots, H_{(m_0)}, H_{(m_0+1)}, H_{(m_0+2)}, \dots, H_{(m_0+j_0)}$ ditolak.

Misalkan $\frac{m_0 + j}{m + 1} q^* = p''$, maka

$$E\left[Q \mid P_{m_0+1} = p_1, \dots, P_m = p_{m_1}\right] = \int_0^{p''} E\left[Q \mid P_{(m_0)}' = p, P_{m_0+1} = p_1, \dots, P_m = p_{m_1}\right] f_{P_{(m_0)}'}(p) dp \\ + \int_{p''}^1 E\left[Q \mid P_{(m_0)}' = p, P_{m_0+1} = p_1, \dots, P_m = p_{m_1}\right] f_{P_{(m_0)}'}(p) dp \quad (5)$$

pada integral bagian pertama, $0 < p < p''$, berdasarkan pertidaksamaan (4)

$p_{j_0} \leq p''$ maka Hipotesis null $H_1, \dots, H_{m_0}, H_{m_0+1}, H_{m_0+2}, \dots, H_{m_0+j_0}$ ditolak

atau sebanyak (m_0+j_0) hipotesis ditolak. Dan $Q = \frac{m_0}{m_0 + j_0}$,

$$E\left[Q \mid P_{(m_0)}' = p, P_{m_0+1} = p_1, \dots, P_m = p_{m_1}\right] = \frac{m_0}{m_0 + j_0}$$

sehingga pada integral bagian pertama,

$$\int_0^{p''} \frac{m_0}{m_0 + j_0} m_0 p^{(m_0-1)} dp = \frac{m_0}{m_0 + j_0} \frac{m_0}{m_0 - 1 + 1} p^{m_0-1+1} \Big|_0^{p''} \\ = \frac{m_0}{m_0 + j_0} p^{m_0} \Big|_0^{p''} \\ = \frac{m_0}{m_0 + j_0} (p'')^{m_0}$$

Gunakan pertidaksamaan (4),

$$\frac{m_0}{m_0 + j_0} \frac{p''}{p''} (p'')^{m_0} = \frac{m_0}{m_0 + j_0} \cdot \frac{m_0 + j_0}{m + 1} \frac{q^*}{p''} (p'')^{m_0} = \frac{m_0}{m_0 + j_0} \cdot \frac{m_0 + j_0}{m + 1} q^* (p'')^{m_0-1} \\ = \frac{m_0}{m + 1} q^* (p'')^{m_0-1} \quad (6)$$

Pada integral bagian kedua, $p'' < p < 1$, berdasarkan pertidaksamaan (4), Karena j_0 dan p'' terdefinisi, maka tidak ada hipotesis yang ditolak, yaitu hipotesis dengan p -value $p_{j+1}, p_{j+2}, \dots, p_{m_1}$, karena yang ditolak hanya hipotesis dengan p -value $p_1, p_2, \dots, p_j = p_{j_0}$.

Oleh karena itu, ketika semua hipotesis (H_0 benar dan H_0 salah) dipertimbangkan bersama dan semua p -value diurutkan. Hipotesis $H_{(i)}$; $i=1, 2, \dots, m$ dapat ditolak hanya jika terdapat k , dimana $i \leq k \leq m_0 + j - 1$ yang memenuhi $p_{(k)} \leq \frac{k}{m+1} q^*$.

$$p_{(k)} \leq \frac{k}{m+1} q^* \text{ ekuivalen dengan } \frac{p_{(k)}}{p} \leq \frac{k}{m_0 + j - 1} \cdot \frac{m_0 + j - 1}{(m+1)p} q^* \quad (7)$$

Ketika bersyarat $P'_{(m_0)} = p, \frac{P'_{(i)}}{p}, i=1, 2, \dots, m_0 - 1$ adalah variabel random

independen berdistribusi $U(0,1)$ dan $\frac{P_i}{p}, i=1, 2, \dots, j$

Dengan menggunakan pertidaksamaan (7) untuk menguji $m_0 + j - 1 = m' \leq m$ hipotesis, ekuivalen menggunakan prosedur (1) dengan konstanta

$\frac{m_0 + j - 1}{(m+1)p} q^*$. Sehingga

$$E \left[Q \mid P'_{(m_0)} = p, P_{m_0+1} = p_1, \dots, P_m = p_{m_1} \right] \leq \frac{m_0 - 1}{m_0 + j - 1} \cdot \frac{m_0 + j - 1}{(m+1)p} q^* = \frac{m_0 - 1}{(m+1)p} q^* \quad (8)$$

Sehingga integral bagian kedua :

$$\begin{aligned}
\int_{p''}^1 \frac{m_0 - 1}{(m + 1)p} q^* m_0 p^{(m_0 - 1)} dp &= \int_{p''}^1 \frac{m_0 - 1}{(m + 1)p} q^* m_0 \frac{p^{m_0}}{p} dp = \int_{p''}^1 \frac{m_0 - 1}{(m + 1)} q^* m_0 \frac{p^{m_0}}{p^2} dp \\
&= \frac{m_0}{m + 1} q^* \int_{p''}^1 (m_0 - 1) p^{(m_0 - 2)} dp \\
&= \frac{m_0}{m + 1} q^* \frac{m_0 - 1}{m_0 - 2 + 1} p^{m_0 - 2 + 1} \Big|_{p''}^1 \\
&= \frac{m_0}{m + 1} q^* \frac{m_0 - 1}{m_0 - 1} p^{m_0 - 1} \Big|_{p''}^1 \\
&= \frac{m_0}{m + 1} q^* (1^{m_0 - 1} - p''^{(m_0 - 1)}) \\
&= \frac{m_0}{m + 1} q^* (1 - p''^{(m_0 - 1)})
\end{aligned} \tag{9}$$

Tambahkan pertidaksamaan (6) dan (9)

$$\begin{aligned}
&\frac{m_0}{m + 1} q^* (p'')^{m_0 - 1} + \frac{m_0}{m + 1} q^* (1 - p''^{(m_0 - 1)}) \\
&= \frac{m_0}{m + 1} q^* (p'')^{m_0 - 1} + \frac{m_0}{m + 1} q^* - \frac{m_0}{m + 1} q^* (p'')^{m_0 - 1} \\
&= \frac{m_0}{m + 1} q^*
\end{aligned}$$

Terbukti bahwa ketika $m_0 < m$,

$$E\left[Q \mid P_{m_0 + 1} = p_1, P_{m_0 + 2} = p_2, \dots, P_{m_0 + m_1} = p_{m_1}\right] \leq \frac{m_0}{m + 1} q^* \text{ benar.}$$

Sehingga untuk $m + 1$ lemma (2) benar.

$$\text{Jadi, } \forall m \text{ berlaku } E\left[Q \mid P_{m_0 + 1} = p_1, P_{m_0 + 2} = p_2, \dots, P_{m_0 + m_1} = p_{m_1}\right] \leq \frac{m_0}{m} q^*$$

Dengan prosedur (1), tingkat signifikansi untuk masing-masing pengujian dipilih $\frac{k}{m}\alpha$; $1 \leq k \leq m$ sedemikian sehingga $FDR \leq \alpha$, sehingga prosedur Benjamin-Hochberg mempunyai power untuk masing-masing pengujian yang lebih baik dibandingkan dengan metode Bonferroni. Akan tetapi, prosedur tersebut ternyata mengontrol FDR pada tingkat signifikansi $\frac{m_0}{m}\alpha$. Prosedur ini menjamin bahwa $FDR \leq \alpha$ dengan mengabaikan banyaknya H_0 benar dan mengabaikan distribusi dari p -value untuk hipotesis H_0 yang salah. Hal inilah yang menjadi kelemahan dari metode Benjamin dan Hochberg.

3.4 POSITIF FALSE DISCOVERY RATE (pFDR)

Pada pengujian hipotesis berganda, terkadang ditemukan beberapa kasus yang menyertakan banyak sekali pengujian dan dapat dipastikan bahwa akan selalu ada hipotesis yang ditolak. Oleh sebab itu, $\Pr(R > 0)$ bernilai satu. Jadi, definisi $FDR = E\left[\frac{V}{R} \mid R > 0\right] \cdot \Pr(R > 0)$ dirubah menjadi

$E\left[\frac{V}{R} \mid R > 0\right] \cdot 1 = E\left[\frac{V}{R} \mid R > 0\right]$, yang dinamakan pFDR (*positif False Discovery Rate*).

Definisi : $pFDR = E\left[\frac{V}{R} \mid R > 0\right]$

3.4.1 Perbandingan Antara FDR dan pFDR

Berikut akan dijelaskan tentang perbandingan antara FDR dan pFDR :

Secara numerik pFDR dekat dengan FDR tetapi secara konsep pFDR mempunyai beberapa keuntungan. Pada FDR, ketika daerah penolakan sangat kecil karena α mendekati 0, maka $\Pr(R>0)$ mendekati 0 dan FDR akan mendekati 0. Sedangkan pada pFDR, dengan kondisi yang sama, bukan berarti kesalahan penolakan akan menurun menuju 0, tetapi kesalahan penolakan akan mendekati nilai $\pi_0 = \frac{m_0}{m}$ yaitu proporsi H_0 benar diantara semua pengujian.

Ketika daerah penolakan sangat kecil hingga mungkin hanya satu *p-value* yang jatuh pada daerah penolakan, maka π_0 digunakan untuk menaksir kesalahan penolakan.

3.4.2 Metode Storey

Storey (2002) memperkenalkan suatu metoda yang bertujuan memperbaiki metoda Benjamin-Hochberg. Pada metoda ini, akan ditetapkan $\pi_0 = \frac{m_0}{m}$, proporsi H_0 benar diantara semua pengujian. Dengan ditetapkannya π_0 , diharapkan pengujian akan mempunyai power yang lebih baik dibandingkan dengan metoda Benjamin-Hochberg.

Berbeda dengan dua metode sebelumnya, pFDR tidak dapat dikontrol apabila ditentukan tingkat signifikansi tertentu. Misalkan pada pengujian hipotesis berganda dengan tingkat signifikansi α , dilakukan m pengujian.

Ketika $m=m_0$, artinya semua H_0 benar, $E\left[\frac{V}{R} \mid R > 0\right] = 1 > \alpha$, sehingga pFDR

tidak dapat dikontrol. Oleh karena itu, Metode Storey yang mengontrol pFDR, diawali dengan menentukan daerah penolakan terlebih dahulu baru kemudian akan menaksir tingkat signifikansi untuk pengujian hipotesis berganda

3.4.3 Interpretasi Bayesian dari pFDR

Misalkan dilakukan m pengujian hipotesis (H_1, H_2, \dots, H_m) dengan statistik uji yang independent T_1, T_2, \dots, T_m . (T_i, H_i) adalah pasangan terurut yang menyatakan bahwa statistik uji ke i adalah statistik uji untuk hipotesis ke i ; $i=1,2,\dots,m$. Diberikan daerah penolakan yang sama, Γ . Definisi positif

false discovery rate :
$$\text{pFDR}(\Gamma) = E\left[\frac{V(\Gamma)}{R(\Gamma)} \mid R(\Gamma) > 0\right]$$

dimana $V(\Gamma) = \#\{\text{null } T_i : T_i \in \Gamma\}$ dan $R(\Gamma) = \#\{T_i : T_i \in \Gamma\}$. V adalah

banyaknya kesalahan tipe 1 atau menolak H_0 ketika H_0 benar, dan R adalah

banyaknya H_0 ditolak. Misalkan, ketika H_0 hipotesis ke i benar dinotasikan

sebagai $H_i=0$ dan ketika H_0 hipotesis ke i salah dinotasikan sebagai $H_i=1$,

untuk $i=1,2,\dots,m$. Misalkan π_0 adalah probabilitas prior bahwa H_0 benar.

Diasumsikan bahwa H_i adalah i.i.d, variabel random Bernoulli dengan

$$\Pr(H_i = 0) = \pi_0 \text{ dan } \Pr(H_i = 1) = 1 - \pi_0 = \pi_1.$$

Ketika $m=1$, terlihat jelas bahwa probabilitas terjadinya suatu kesalahan ketika hipotesis ditolak adalah probabilitas H_0 benar ketika H_0 ditolak yaitu $\Pr(H = 0 | T \in \Gamma)$. Akan ditunjukkan bahwa ketika $m > 1$, probabilitas terjadinya suatu kesalahan ketika hipotesis ditolak adalah $\Pr(H = 0 | T \in \Gamma)$ juga.

Akan ditunjukkan $\text{pFDR} = \Pr(H = 0 | T \in \Gamma)$:

$$\begin{aligned} \text{pFDR}(\Gamma) &= E\left(\frac{V}{R} | R > 0\right) = E\left[E\left[\left(\frac{V}{R} | R > 0\right) | R = i\right]\right] \\ &= \sum_{i=1}^m E\left[\left(\frac{V}{R} | R > 0\right) | R = i\right] \Pr(R = i) \\ &= \sum_{i=1}^m E\left(\frac{V}{i} | R = i\right) \Pr(R = i | R > 0) \end{aligned}$$

Misalkan didefinisikan,

$$\text{fungsi operator : } \begin{cases} 1 & \text{jika } T_j \in \Gamma \\ 0 & \text{jika } T_j \notin \Gamma \end{cases}$$

$$\text{dan fungsi operator: } \begin{cases} 1 & \text{jika } H_j = 0 \\ 0 & \text{jika } H_j = 1 \end{cases}$$

Terdapat empat kemungkinan yang terjadi, yaitu :

- $1(T_j \in \Gamma)1(H_j = 0) = H_0$ ditolak ketika H_0 benar
- $1(T_j \in \Gamma)0(H_j = 1) = H_0$ ditolak ketika H_0 salah
- $0(T_j \notin \Gamma)1(H_j = 0) = H_0$ tidak ditolak ketika H_0 benar

- $0(T_j \notin \Gamma)0(H_j = 1) = H_0$ tidak ditolak ketika H_0 salah

Maka, $V = \sum_{i=1}^m 1(T_i \in \Gamma)1(H_i = 0)$ dan

$$E(V|R=i) = E\left(\sum_{j=1}^m 1(T_j \in \Gamma)1(H_j = 0)|R=i\right)$$

karena statistik uji adalah independen dan $R=i$ berarti $T_1, T_2, \dots, T_i \in \Gamma$ dan

$T_{i+1}, T_{i+2}, \dots, T_m \notin \Gamma$, sehingga

$$E(V|R=i) = E\left(\sum_{j=1}^m 1(T_j \in \Gamma)1(H_j = 0)|T_1, \dots, T_i \in \Gamma, T_{i+1}, \dots, T_m \notin \Gamma\right)$$

Untuk $j=1, 2, \dots, i$ sudah pasti $T_j \in \Gamma$ sehingga,

$$\begin{aligned} E(V|R=i) &= E\left(\sum_{j=1}^i 1(H_j = 0)|T_1, \dots, T_i \in \Gamma, T_{i+1}, \dots, T_m \notin \Gamma\right) \\ &= \sum_{j=1}^i E(1(H_j = 0)|T_j \in \Gamma) \\ &= E(1(H_1 = 0)|T_1 \in \Gamma) + E(1(H_2 = 0)|T_2 \in \Gamma) + \dots + E(1(H_i = 0)|T_i \in \Gamma) \\ &= 1 \cdot \Pr(H_1 = 0|T_1 \in \Gamma) + 1 \cdot \Pr(H_2 = 0|T_2 \in \Gamma) + \dots + 1 \cdot \Pr(H_i = 0|T_i \in \Gamma) \\ &= \Pr(H_1 = 0|T_1 \in \Gamma) + \Pr(H_2 = 0|T_2 \in \Gamma) + \dots + \Pr(H_i = 0|T_i \in \Gamma) \end{aligned}$$

karena ketika $m=1$, terlihat jelas bahwa probabilitas terjadinya suatu

kesalahan ketika hipotesis ditolak yaitu $\Pr(H = 0|T \in \Gamma)$ maka

$\Pr(H_j = 0|T_j \in \Gamma)$; $j = 1, 2, \dots, i$ nilainya sama, yaitu $\Pr(H = 0|T \in \Gamma)$.

Sehingga $E(V|R=i) = i \Pr(H = 0|T \in \Gamma)$

jadi,

$$\begin{aligned}
\text{pFDR}(\Gamma) &= \sum_{i=1}^m E\left(\frac{V}{i} | R=i\right) \Pr(R=i | R>0) \\
&= \sum_{i=1}^m \frac{i \Pr(H=0 | T \in \Gamma)}{i} \cdot \Pr(R=i | R>0) \\
&= \sum_{i=1}^m \Pr(H=0 | T \in \Gamma) \Pr(R=i | R>0) \\
&= \Pr(H=0 | T \in \Gamma) \sum_{i=1}^m \Pr(R=i | R>0) \\
&= \Pr(H=0 | T \in \Gamma) \cdot 1 \\
&= \Pr(H=0 | T \in \Gamma)
\end{aligned}$$

pFDR dapat ditulis sebagai probabilitas posterior:

$$\begin{aligned}
\text{pFDR}(\Gamma) &= \Pr(H=0 | T \in \Gamma) \\
&= \frac{\pi_0 \Pr(T \in \Gamma | H=0)}{\Pr(T \in \Gamma)}
\end{aligned}$$

dimana $\Pr(T \in \Gamma) = \pi_0 \Pr(T \in \Gamma | H=0) + \pi_1 \Pr(T \in \Gamma | H=1)$

sehingga

$$\begin{aligned}
\text{pFDR}(\Gamma) &= \Pr(H=0 | T \in \Gamma) \\
&= \frac{\pi_0 \Pr(T \in \Gamma | H=0)}{\pi_0 \Pr(T \in \Gamma | H=0) + \pi_1 \Pr(T \in \Gamma | H=1)} \\
&= \frac{\pi_0 \cdot \{\text{kesalahan tipe 1 dari } \Gamma\}}{\pi_0 \cdot \{\text{kesalahan tipe 1 dari } \Gamma\} + \pi_1 \cdot \{\text{power dari } \Gamma\}}
\end{aligned}$$

Ini menunjukkan bahwa pFDR meningkat dengan meningkatnya kesalahan tipe1 dan pFDR menurun dengan meningkatnya power.

3.4.4 Menaksir positif False Discovery Rate

Misalkan dilakukan m pengujian hipotesis identik dan daerah penolakan yang sama untuk setiap pengujian yaitu Γ . Untuk daerah penolakan berdasarkan p -value, semua daerah penolakan dibentuk $[0, \gamma]$ untuk $\gamma \geq 0$, sehingga daerah penolakan Γ dirubah menjadi $[0, \gamma]$. Sehingga teorema ditulis sebagai :

$$pFDR(\gamma) = \frac{\Pr(T \in \Gamma | H = 0)}{\Pr(T \in \Gamma)} = \frac{\pi_0 \Pr(P \leq \gamma | H = 0)}{\Pr(P \leq \gamma)} = \frac{\pi_0 \gamma}{\Pr(P \leq \gamma)}$$

Dimana P menyatakan p -value dari pengujian dan π_0 adalah probabilitas prior. Dalam tugas akhir ini π_0 ditetapkan terlebih dahulu.

$$\widehat{\Pr}(P \leq \gamma) = \frac{\#\{p_i \leq \gamma\}}{m} = \frac{R(\gamma) \vee 1}{m}$$

$$\text{Sehingga, } \widehat{pFDR}(\gamma) = \frac{\pi_0 \gamma}{\widehat{\Pr}(P \leq \gamma)} = \frac{\pi_0 \gamma}{\frac{R(\gamma) \vee 1}{m}} = \frac{\pi_0 m \gamma}{R(\gamma) \vee 1}.$$

Pada $pFDR$, bersyarat $R(\gamma) > 0$ dan $1 - (1 - \gamma)^m$ adalah batas bawah untuk

$$\Pr\{R(\gamma) > 0\}. \text{ Oleh karena itu, } \widehat{pFDR}(\gamma) = \frac{\pi_0 m \gamma}{\{R(\gamma) \vee 1\} \{1 - (1 - \gamma)^m\}}.$$

Sedangkan pada FDR , tidak bersyarat $R(\gamma) > 0$, maka dengan metode

$$\text{Storey, dapat pula dicari taksiran dari } FDR, \text{ yaitu } \widehat{FDR}(\gamma) = \frac{\pi_0 m \gamma}{R(\gamma) \vee 1}.$$

3.4.5 *q-value*

q value similar dengan *p value*, *q value* dari suatu pengujian mengukur infimum pFDR yang terjadi ketika melakukan uji signifikansi. *p value* biasa digunakan untuk melakukan pengujian hipotesis tunggal sedangkan *q value* berguna untuk melakukan pengukuran signifikansi dari setiap uji dari banyak pengujian yang dilakukan secara simultan.

Definisi 1

Untuk sebuah nilai statistik uji $T=t$, *q-value* dari t didefinisikan:

$$q(t) = \inf_{\{\Gamma: t \in T\}} \{pFDR(\Gamma)\}$$

Definisi 2

q-value dari observed *p-value* adalah

$$q(p) = \inf_{\gamma \geq p} \{pFDR(\gamma)\} = \inf_{\gamma \geq p} \left\{ \frac{\pi_0 \gamma}{\Pr(P \leq \gamma)} \right\}$$

Algoritma menghitung *q-value* :

1. Untuk m pengujian hipotesis, hitung *p-value* : p_1, p_2, \dots, p_m
2. Urutkan *p-value* : $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(m)}$
3. Hitung : $\hat{q}(p_{(m)}) = \widehat{pFDR}(p_{(m)})$
4. Hitung : $\hat{q}(p_{(i)}) = \min \{pFDR(p_{(i)}), \hat{q}(p_{(i+1)})\}$; $i = m-1, m-2, \dots, 1$

BAB IV
APLIKASI PENGUJIAN HIPOTESIS BERGANDA PADA DATA DNA
MICROARRAY EXPERIMENT

Untuk melengkapi pembahasan mengenai pengujian hipotesis berganda, pada bab ini akan membahas contoh kasus yang dapat diselesaikan dengan metode –metode pengujian untuk hipotesis berganda.

4.1 LATAR BELAKANG MASALAH

Menurut WHO, 8-9% wanita akan mengalami kanker payudara. Ini menjadikan kanker payudara sebagai jenis kanker yang paling banyak ditemui pada wanita. Setiap tahun lebih dari 250.000 kasus baru kanker payudara terdiagnosa di Eropa dan kurang lebih 175.000 di Amerika Serikat. Masih menurut WHO, tahun 2000 diperkirakan 1,2 juta wanita terdiagnosis kanker payudara dan lebih dari 700.000 meninggal karenanya. Belum ada data statistik yang akurat di Indonesia, namun data yang terkumpul dari rumah sakit menunjukkan bahwa kanker payudara menduduki ranking pertama diantara kanker lainnya pada wanita. Kanker payudara merupakan penyebab utama kematian pada wanita akibat kanker.

Penyebab pasti kanker payudara tidak diketahui. Meskipun demikian, riset mengidentifikasi sejumlah faktor yang dapat meningkatkan risiko pada individu tertentu, yang meliputi:

- Keluarga yang memiliki riwayat penyakit serupa
- Usia yang makin bertambah
- Tidak memiliki anak / tidak menyusui
- Kehamilan pertama pada usia di atas 30 tahun
- Periode menstruasi yang lebih lama (menstruasi pertama lebih awal atau menopause lebih lambat)
- Faktor hormonal (baik estrogen maupun androgen).

Dari faktor risiko tersebut di atas, riwayat keluarga serta usia menjadi faktor terpenting. Riwayat keluarga yang pernah mengalami kanker payudara meningkatkan resiko berkembangnya penyakit ini. Wanita yang memiliki sejarah kesehatan keluarganya mengidap kanker payudara kemungkinan besar mereka memiliki sejumlah mutasi kanker payudara 1 (BRCA 1) atau kanker payudara 2 (BRCA2). Para peneliti juga menemukan bahwa kerusakan dua gen yaitu BRCA1 atau BRCA2 dapat meningkatkan risiko wanita terkena kanker payudara 50-85%. Para peneliti menggunakan pembuktian amplifikasi dan berdasarkan metoda DNA dan RNA untuk mendeteksi adanya perubahan susunan gen (mutasi genetik) pada BRCA1 dan BRCA2 melalui penggunaan teknologi microarray DNA.

Pengembangan teknologi microrarray DNA ini sangat bermanfaat dalam menghasilkan sejumlah besar data ekspresi gen. Profiling ekspresi

gen melalui microarray tersebut adalah teknik yang sangat efektif dalam usaha menghimpun ribuan tingkatan ekspresi gen secara simultan. Eksperimen yang di dasarkan ekspresi gen ini dapat di lakukan dengan dua cara: (1) Pengamatan setiap gen dari beberapa kondisi yang berbeda atau (2) Mengevaluasi setiap gen dari satu kondisi tetapi dalam tipe jaringan yang berbeda, khususnya jaringan sel yang mengandung kanker (Furey, et. al. 2000)

4.2 PERMASALAHAN

Penyakit kanker payudara karena keturunan disebabkan oleh mutasi kanker BRCA1 atau BRCA2. Sampel RNA diambil dari 7 pasien pembawa mutasi kanker BRCA1 dan 7 pasien pembawa mutasi kanker BRCA2. Dalam eksperimen DNA microarray yang melibatkan 100 gen, ingin di deteksi gen-gen yang menunjukkan perbedaan ekspresi diantara 2 mutasi kanker BRCA1 dan BRCA2

4.3. SUMBER DATA

Dari *National Human Genome Research Institute, National Institute Of Health, Bethesda.; the Departement of Oncology. University of Lund, Lund, Sweden.*

4.4 ANALISIS DATA

Tujuan : Bagaimana menentukan pengaruh atau efek dari variabel independen “mutasi kanker ” terhadap variabel dependen “nilai ekspresi gen”.

Pertama-tama dilakukan uji secara keseluruhan yaitu menguji apakah ada pengaruh variabel independen mutasi kanker terhadap ke seratus variabel gen secara sekaligus.

- Hipotesis :

$$H_0 : \begin{pmatrix} \mu_{11} \\ \mu_{21} \\ \vdots \\ \mu_{1001} \end{pmatrix} = \begin{pmatrix} \mu_{12} \\ \mu_{22} \\ \vdots \\ \mu_{1002} \end{pmatrix},$$

H_1 : tidak demikian

dimana :

μ_{11} : mean nilai ekspresi gen1 untuk BRCA1

μ_{21} : mean nilai ekspresi gen2 untuk BRCA1

\vdots

μ_{1002} : mean nilai ekspresi gen100 untuk BRCA2

- Tingkat signifikansi = 0.05
- Aturan keputusan : H_0 ditolak jika $\hat{\alpha} < \alpha = 0.05$

Dengan menggunakan software SPSS 16 diperoleh output *Multivariate test* yang dapat dilihat pada lampiran 1:

- Keputusan : karena $\hat{\alpha} = 0.646 > \alpha = 0.05$, maka H_0 tidak ditolak.

- Kesimpulan : Dengan tingkat kepercayaan 95%, dapat disimpulkan bahwa tidak ada pengaruh mutasi kanker terhadap variabel nilai ekspresi gen 1,2,...,100

Akan tetapi, ingin dilihat apakah ada pengaruh variabel independen mutasi kanker terhadap masing-masing ke seratus variabel dependen nilai ekspresi gen. Kemudian, dilakukan masing-masing 100 pengujian pada satu waktu secara bersama-sama.

- Hipotesis :

$$H_0 : \mu_{i1} = \mu_{i2} \quad ; i = 1, 2, \dots, 100$$

H_1 : tidak demikian

Dengan menggunakan software SPSS 16, diperoleh output *test between subject* yang dapat dilihat pada lampiran 2. Kemudian, dengan menggunakan software *q-value*, dihasilkan *q-value* untuk masing-masing pengujian. Diberikan tabel untuk $\hat{\alpha}$ atau *p-value* yang telah diurutkan dari yang terkecil ke terbesar dan *q-value* untuk masing-masing pengujian, dapat dilihat pada lampiran 3.

- Tingkat signifikansi dan Aturan keputusan :

1. Metode Bonferroni :

Tingkat signifikansi untuk masing-masing pengujian dipilih $\frac{\alpha}{m}$.

Aturan keputusan : H_0 ditolak jika $\hat{\alpha} \leq \frac{\alpha}{m} = \frac{0.05}{100} = 0.0005$

2. Metode Benjamin-Hochberg :

Tingkat signifikansi untuk masing-masing pengujian dipilih $\frac{k}{m}\alpha$.

Dengan mengikuti aturan berikut ini :

➤ Urutkan *p-value* $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(100)}$.

➤ Hitung $\hat{k} = \max_{1 \leq k \leq m} \left\{ k : p_{(k)} \leq \frac{k}{m}\alpha \right\}$. maka H_0 dengan *p-value*

$p_{(1)}, p_{(2)}, \dots, p_{(k)}$ ditolak.

3. Metode Storey :

Tingkat signifikansi untuk masing-masing pengujian dipilih α .

Aturan keputusan : H_0 ditolak jika *q-value* $< \alpha = 0.05$

• Keputusan :

1. Dengan menggunakan Metode Bonferroni, dari tabel *test between subject* pada lampiran 2, diperoleh bahwa tidak ada nilai

$\hat{\alpha} < \frac{\alpha}{m} = \frac{0.05}{100} = 0.0005$, sehingga tidak ada H_0 yang ditolak. Artinya

tidak ada pengaruh mutasi kanker terhadap nilai ekspresi gen.

2. Dengan metode Benjamin-Hochberg, dari tabel pada lampiran 3, diperoleh bahwa tidak ada nilai *p-value* yang memenuhi

$\hat{k} = \max_{1 \leq k \leq m} \left\{ k : p_{(k)} \leq \frac{k}{m}\alpha \right\}$, sehingga tidak ada H_0 yang ditolak. Artinya

tidak ada pengaruh mutasi kanker terhadap nilai ekspresi gen.

3. Dengan menggunakan metode Storey, dipilih $\pi_0 = 0.614$, dari tabel pada lampiran 3, diperoleh bahwa pada gen ke 10 memiliki q-value = $0.0457 < 0.05$ maka H_0 ditolak, artinya ada pengaruh mutasi kanker terhadap nilai ekspresi gen ke 10



BAB V

KESIMPULAN DAN SARAN

5.1 KESIMPULAN

Dari hasil yang diperoleh dalam penulisan skripsi ini, dapat disimpulkan bahwa :

1. Metode – metode pengujian untuk hipotesis berganda digunakan untuk mengontrol kesalahan tipe 1 secara keseluruhan dalam *family hipotesis*. Ada beberapa cara untuk mengukur kesalahan tipe 1 secara keseluruhan pada pengujian hipotesis berganda, diantaranya adalah *Family Wise Error Rate* (FWER), *False Discovery Rate* (FDR), dan *positif False Discovery Rate* (pFDR).
2. Metode pengujian untuk hipotesis berganda diantaranya adalah Metode Bonferroni salah satu metode untuk FWER, metode Benjamin-Hochberg untuk FDR yang memperbaiki metode Bonferroni, dan metode Storey untuk pFDR yang memperbaiki metode Storey.
3. Pada pengujian hipotesis berganda untuk kasus mendeteksi perbedaan ekspresi gen diantara dua mutasi kanker, dengan dua metode sebelumnya tidak ditemukan adanya perbedaan ekspresi gen , namun dengan metode Storey dapat ditunjukkan adanya perbedaan ekspresi

gen. Jadi, dengan metode storey dapat dipastikan bahwa akan selalu ada hipotesis yang ditolak bila dilakukan banyak pengujian.

5.2 SARAN

1. Agar diperoleh nilai π_0 yang menyatakan proporsi H_0 benar di antara semua pengujian sesuai dengan kondisi sebenarnya, disarankan untuk menaksir π_0
2. Ada beberapa metode yang disarankan untuk menaksir π_0 , antara lain metode Bootstrap, *polynomial Bernstein*.
3. Pembahasan untuk mengukur kesalahan tipe 1 pada pengujian hipotesis berganda dapat dikembangkan untuk kasus hipotesis-hipotesis dependen.

DAFTAR PUSTAKA

- Walpole, Ronald W. (1995): *Pengantar Statistika*. 3th ed. PT Gramedia Pustaka Utama, Jakarta.
- Ross, Sheldon. (2002): *A First Course in Probability*. 6th ed. Prentice Hall, Upper Saddle River, New Jersey.
- Hogg, Robert V., Joseph W. McKean, Allen T. Craig. (2005): *Introduction to Mathematical Statistics*. 6th ed. Pearson Prentice Hall, United States of America.
- Benjamini, Yoav. and Hochberg, Yocef. (1995): *Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing*, *Journal of the Royal Statistic of Society*, 289-300.
- Storey, John D. (2002): *A Direct Approach to False Discovery Rates*, *Journal of the Royal Statistic of Society*, 479 – 498.
- Storey, John D. (2003): *The Positive False Discovery Rate: A Bayesian Interpretation and the q-value*, *Journal of the Annals of Statistics*, 2013 – 2035.
- Storey, John D. (2003): *Statistical Significance for Genomewide Studies*, *Journal of Proceedings of the National Academy of Sciences the United States of America*, 9440 – 9445.

LAMPIRAN 1

Output *Multivariate Test*

Multivariate Tests^b

Effect		Value	F	Hypothesis df	Error df	Sig.
Intercept	Pillai's Trace	.999	1.256E2 ^a	12.000	1.000	.070
	Wilks' Lambda	.001	1.256E2 ^a	12.000	1.000	.070
	Hotelling's Trace	1.507E3	1.256E2 ^a	12.000	1.000	.070
	Roy's Largest Root	1.507E3	1.256E2 ^a	12.000	1.000	.070
BRCA	Pillai's Trace	.928	1.074 ^a	12.000	1.000	.646
	Wilks' Lambda	.072	1.074 ^a	12.000	1.000	.646
	Hotelling's Trace	12.889	1.074 ^a	12.000	1.000	.646
	Roy's Largest Root	12.889	1.074 ^a	12.000	1.000	.646

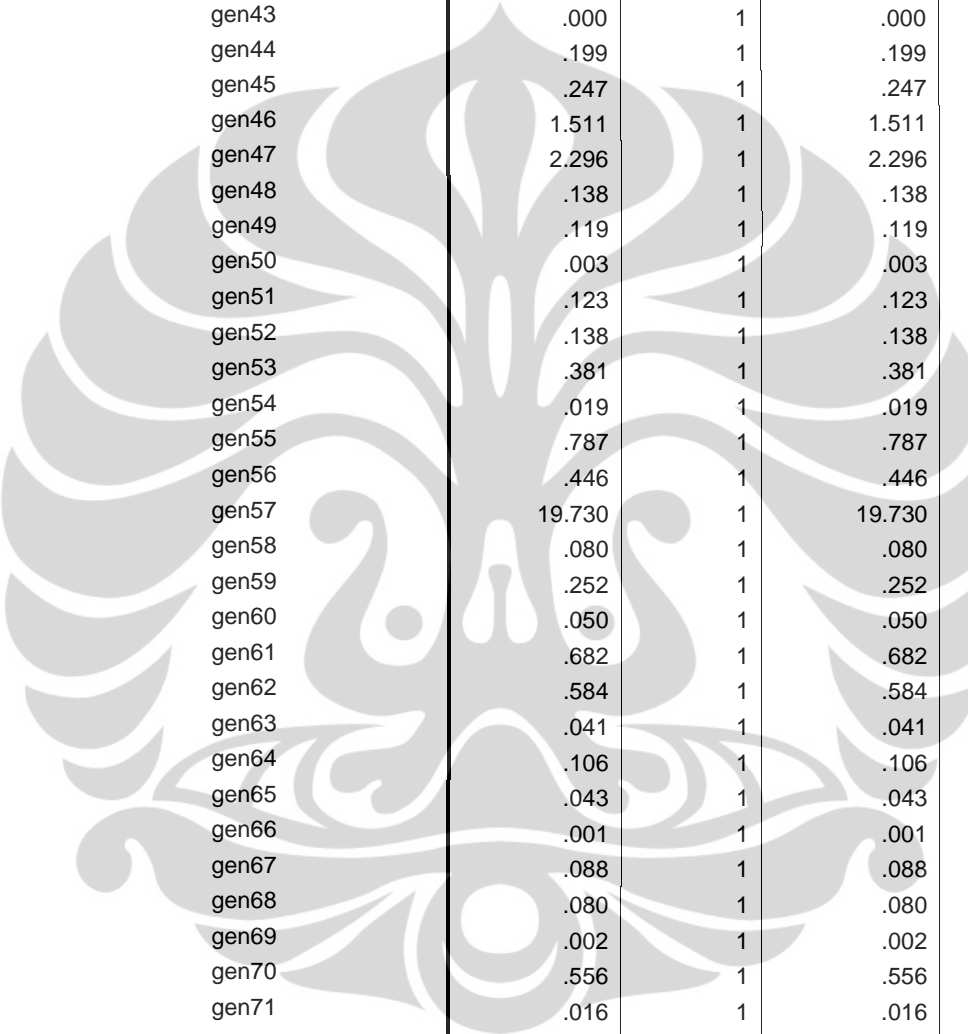
a. Exact statistic

b. Design: Intercept + BRCA

LAMPIRAN 2

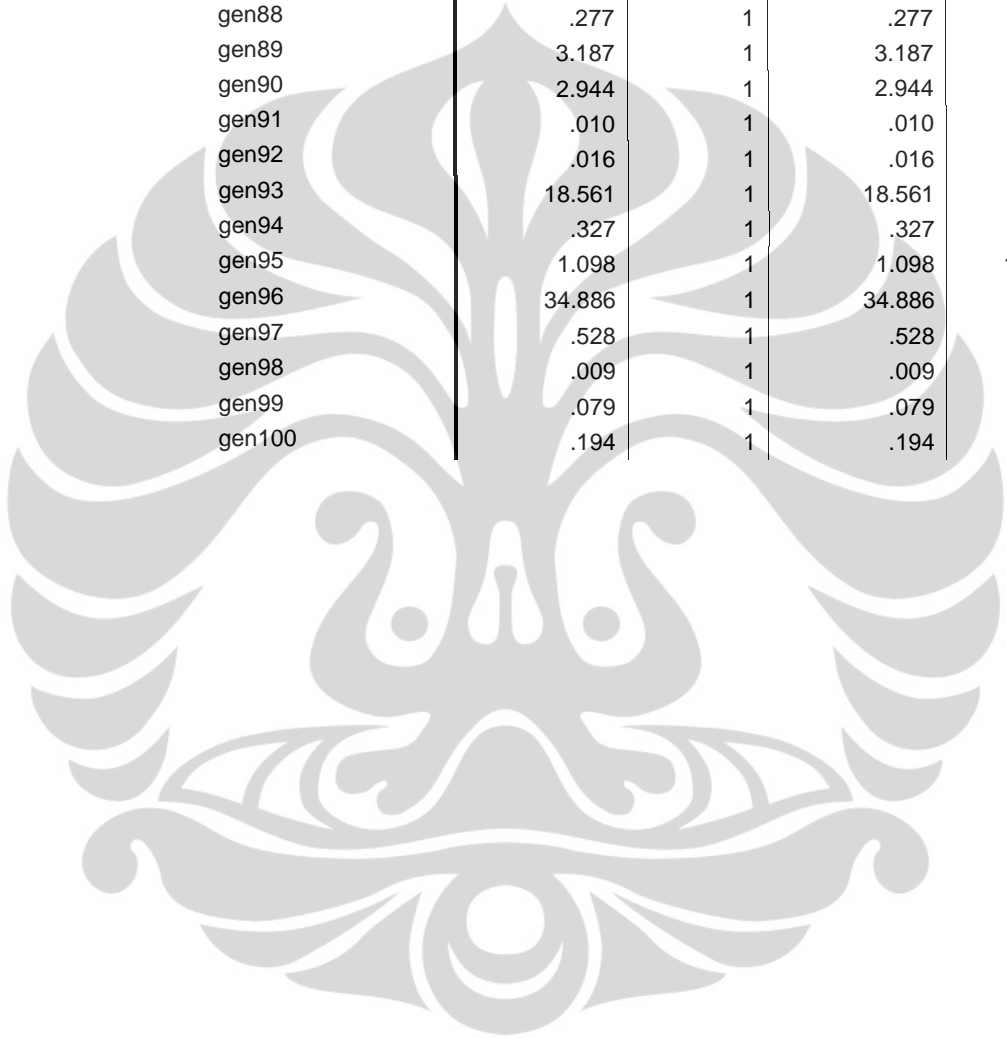
Output Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
BRCA	gen1	.704	1	.704	5.304	.040
	gen2	.425	1	.425	5.029	.045
	gen3	.187	1	.187	.950	.349
	gen4	2.200	1	2.200	3.652	.080
	gen5	.126	1	.126	.278	.607
	gen6	.101	1	.101	1.313	.274
	gen7	.007	1	.007	.075	.788
	gen8	.119	1	.119	3.448	.088
	gen9	.498	1	.498	3.581	.083
	gen10	.811	1	.811	20.115	.001
	gen11	.713	1	.713	2.962	.111
	gen12	.475	1	.475	10.243	.008
	gen13	.552	1	.552	1.466	.249
	gen14	.000	1	.000	.004	.953
	gen15	.221	1	.221	.602	.453
	gen16	.022	1	.022	.235	.636
	gen17	1.010	1	1.010	3.708	.078
	gen18	.745	1	.745	13.385	.003
	gen19	.327	1	.327	2.420	.146
	gen20	1.207	1	1.207	3.904	.072
	gen21	.069	1	.069	.334	.574
	gen22	.058	1	.058	.810	.386
	gen23	1.591	1	1.591	3.092	.104
	gen24	.023	1	.023	.075	.788
	gen25	7.14E-006	1	7.14E-006	.000	.993
	gen26	7.14E-006	1	7.14E-006	.000	.993
	gen27	.172	1	.172	.279	.607
	gen28	.117	1	.117	1.182	.298
	gen29	.368	1	.368	.643	.438
	gen30	1.708	1	1.708	10.671	.007
	gen31	.031	1	.031	.287	.602
	gen32	.006	1	.006	.197	.665
	gen33	.117	1	.117	2.372	.149
	gen34	.003	1	.003	.003	.957
	gen35	.002	1	.002	.012	.915



gen36	2.853	1	2.853	10.099	.008
gen37	.009	1	.009	.300	.594
gen38	.146	1	.146	3.099	.104
gen39	.432	1	.432	2.806	.120
gen40	.358	1	.358	1.621	.227
gen41	.019	1	.019	.095	.763
gen42	.043	1	.043	.147	.709
gen43	.000	1	.000	.004	.950
gen44	.199	1	.199	10.151	.008
gen45	.247	1	.247	10.727	.007
gen46	1.511	1	1.511	4.843	.048
gen47	2.296	1	2.296	2.316	.154
gen48	.138	1	.138	1.127	.309
gen49	.119	1	.119	3.080	.105
gen50	.003	1	.003	.152	.704
gen51	.123	1	.123	1.747	.211
gen52	.138	1	.138	3.794	.075
gen53	.381	1	.381	1.978	.185
gen54	.019	1	.019	.363	.558
gen55	.787	1	.787	.102	.755
gen56	.446	1	.446	.318	.583
gen57	19.730	1	19.730	1.173	.300
gen58	.080	1	.080	1.061	.323
gen59	.252	1	.252	4.918	.047
gen60	.050	1	.050	.914	.358
gen61	.682	1	.682	7.906	.016
gen62	.584	1	.584	1.337	.270
gen63	.041	1	.041	.495	.495
gen64	.106	1	.106	.244	.631
gen65	.043	1	.043	.185	.675
gen66	.001	1	.001	.117	.738
gen67	.088	1	.088	.438	.521
gen68	.080	1	.080	.845	.376
gen69	.002	1	.002	.056	.817
gen70	.556	1	.556	.857	.373
gen71	.016	1	.016	.056	.816
gen72	.001	1	.001	.012	.915
gen73	.019	1	.019	.404	.537
gen74	.415	1	.415	6.374	.027
gen75	.016	1	.016	.364	.557
gen76	.005	1	.005	.016	.903
gen77	.528	1	.528	2.327	.153
gen78	.263	1	.263	1.128	.309
gen79	.027	1	.027	.079	.783
gen80	.544	1	.544	2.551	.136

gen81	.015	1	.015	.946	.350
gen82	11.666	1	11.666	3.222	.098
gen83	.831	1	.831	6.802	.023
gen84	7.726	1	7.726	2.816	.119
gen85	.365	1	.365	1.299	.277
gen86	.124	1	.124	2.594	.133
gen87	.057	1	.057	.313	.586
gen88	.277	1	.277	.456	.512
gen89	3.187	1	3.187	4.693	.051
gen90	2.944	1	2.944	4.994	.045
gen91	.010	1	.010	.133	.721
gen92	.016	1	.016	.546	.474
gen93	18.561	1	18.561	5.030	.045
gen94	.327	1	.327	.643	.438
gen95	1.098	1	1.098	10.192	.008
gen96	34.886	1	34.886	6.231	.028
gen97	.528	1	.528	2.260	.159
gen98	.009	1	.009	.071	.795
gen99	.079	1	.079	1.842	.200
gen100	.194	1	.194	1.083	.319



LAMPIRAN 3

Nilai *p-value* dan *q-value* yang Telah Diurutkan dari yang Terkecil ke yang Terbesar

gen ke	<i>p-value</i>	<i>q-value</i>
10	0.000745931	0.04578698
18	0.003275149	0.06101816
45	0.006637969	0.06101816
30	0.006744529	0.06101816
12	0.007626529	0.06101816
95	0.007740273	0.06101816
44	0.00783186	0.06101816
36	0.007952537	0.06101816
61	0.015696817	0.1070564
83	0.022883838	0.1404663
74	0.02667779	0.1437946
96	0.028111263	0.1437946
1	0.0399754	0.1639205
93	0.044572966	0.1639205
2	0.044600948	0.1639205
90	0.045215484	0.1639205
59	0.046625961	0.1639205
46	0.048068714	0.1639205
89	0.05112784	0.1651761
20	0.071608537	0.2118229
52	0.075216818	0.2118229
17	0.078168994	0.2118229
4	0.080180705	0.2118229
9	0.082821065	0.2118229
8	0.088015943	0.2161049
82	0.09784932	0.2217254
38	0.10379735	0.2217254
23	0.104127157	0.2217254
49	0.104753862	0.2217254
11	0.110886155	0.2268817
84	0.119183618	0.2296997
39	0.11974768	0.2296997
86	0.133233823	0.2459664

80	0.136242112	0.2459664
19	0.145765855	0.2486766
33	0.14949356	0.2486766
77	0.153070927	0.2486766
47	0.1539484	0.2486766
97	0.158639415	0.2496835
53	0.1849667	0.2838422
99	0.199718397	0.2990044
51	0.210844978	0.3081466
40	0.227047295	0.3241091
13	0.24932946	0.3478278
62	0.269993243	0.3612587
6	0.274227062	0.3612587
85	0.276613109	0.3612587
28	0.298294921	0.3722935
57	0.300111952	0.3722935
78	0.309082524	0.3722935
48	0.309323008	0.3722935
100	0.318616666	0.3744343
58	0.323301816	0.3744343
3	0.349043121	0.3904543
81	0.349856162	0.3904543
60	0.35800225	0.3924109
70	0.37280929	0.3981318
68	0.376193724	0.3981318
22	0.38584843	0.4014284
29	0.438203846	0.4410527
94	0.438305516	0.4410527
15	0.452810489	0.4482994
92	0.474173809	0.4619983
63	0.495124371	0.4748732
88	0.512457829	0.4839362
67	0.520813302	0.4843747
73	0.53680213	0.4906363
75	0.557300361	0.4906363
54	0.557834162	0.4906363
21	0.573815378	0.4906363
56	0.583110449	0.4906363
87	0.58639112	0.4906363
37	0.593655995	0.4906363
31	0.601913719	0.4906363
27	0.607087164	0.4906363
5	0.607477077	0.4906363
64	0.630512081	0.5008036
16	0.636383148	0.5008036

32	0.664873304	0.5166009
65	0.675124384	0.5180088
50	0.703737373	0.5303954
42	0.708549516	0.5303954
91	0.721241036	0.533391
66	0.738362305	0.5395524
55	0.755240233	0.542014
41	0.763209086	0.542014
79	0.782842883	0.542014
24	0.788247084	0.542014
7	0.78846757	0.542014
98	0.794711745	0.542014
71	0.816361554	0.5454185
69	0.817474694	0.5454185
76	0.902936518	0.5914121
72	0.914743904	0.5914121
35	0.915314798	0.5914121
43	0.949662407	0.5996775
14	0.952820217	0.5996775
34	0.957415578	0.5996775
25	0.993382673	0.6097614
26	0.993382673	0.6097614