**Analisis Diskriminan Pada Konsumen Air Mineral Dalam Kemasan (AMDK)**

***Discriminant Analysis In Consumer Bottled Of Mineral Water* (*Mineral Water*)**

**Juli Andri1, Rito Goejantoro2, Haeruddin3**

1Mahasiswa Program Studi Statistika FMIPA Universitas Mulawarman

2,3Dosen Program Studi Statistika FMIPA Universitas Mulawarman

***Abstract***

*Discriminant analysis is a method that can be used to determine which variables distinguish between a group with other groups in the population. Constitue one of the methods used in multivariate analysis dependency method, discriminant analysis is used in cases where the independent variable is data metric and the dependent variable non metric form of data. Can also be used to classify data based on different characteristics of data. Therefore, discriminant analysis can be used as a method of grouping. This study aims to determine how the function or discriminant models of consumer behavior on the issue of bottled drinking water and consuming preformance to classify a new object to be included into the category a little or a lot to consume bottled drinking water. Independent variable is the age of the consumer, weight, height, consumer, consumer income, preformance daily campus activities and sports activities consumers. The sample of this study were students force statistics 2012. From the analysis of the model obtained discriminant function is formed based on the results of data processing are:Y* = 7,557 - 0,75 *age* + 0,004 *student allowance* + 0,558 *campus activitie* + 0,301 *sports.* *Based on testing with discriminant method can be concluded that there are only four variables that influence the consumption of bottled drinking water is the variable age, student allowance, the day campus activities and sports events of the week. To classify a new object in the category of a little or a lot can be seen from Z scored the discriminant.*

*Keywords: Bottled drinking water, Discriminant analysis, Discriminant Function Model.*

**Pendahuluan**

Analisis multivariat merupakan objek kajian pada statistika yang mempelajari perilaku dengan hubungan dua atau lebih variabel, maka bisa dikatakan analisis multivariat merupakan perluasan dari analisis univariat atau bivariat. Statistika multivariat pada saat ini diterapkan hampir di setiap cabang ilmu, baik ilmu pengetahuan alam maupun ilmu pengetahuan sosial. Tehnik-tehniknya disukai karena dianggap mampu memodelkan kerumitan sistem yang nyata meskipun sulit diterapkan (Suryanto, 1990).

 Analisis diskriminan merupakan salah satu metode yang digunakan dalam analisis multivariat dengan metode depedensi (dimana hubungan antar variabel sudah bisa dibedakan mana variabel terikat mana variabel bebas). Analisis diskriminan digunakan pada kasus di mana variabel bebas berupa data matrik (interval atau rasio) dan variabel terikat berupa data nonmetrik (nominal atau ordinal) (Johnson dan Wichern, 2002).

Tehnik analisis diskriminan dibedakan menjadi dua bagian, yaitu analisis diskriminan dua kelompok atau kategori dan analisis diskriminan berganda. Dalam analisis diskriminan dua kelompok atau kategori, jika variabel dependen (*Y*) dikelompokan menjadi dua maka dibutuhkan satu fungsi diskriminan. Untuk analisis diskriminan berganda, jika variabel dependen (*Y*) dikelompokan menjadi lebih dari dua kelompok maka diperlukan fungsi diskriminan sebanyak (*k*-1) untuk *k* kategori.

Air adalah kebutuhan dasar (primer) yang tidak dapat dipisahkan dari kehidupan manusia yang menduduki urutan kedua setelah udara. Kebutuhan masyarakat akan air minum layak dan aman dikonsumsi semakin lama kesadaran masyarakat semakin tinggi tentang pentingnya minum air sebagai salah satu kebutuhan esensial untuk beraktivitas dalam kehidupan sehari-hari inilah yang menjadi cikal bakal meningkatnya prospek usaha air minum dalam kemasan (AMDK) yang memasukan produk air minum sehingga menjadi alternatif bagi masyarakat terutama dalam memenuhi kebutuhan air bersih yang layak dan aman untuk dikonsumsi setiap hari, sebagian besar masyarakat indonesia sudah tidak asing lagi dengan AMDK dan mengkonsumsinya untuk kebutuhan sehari-hari, setiap orang harus minum air dalam jumlah yang cukup tidak terkecuali mahasiswa, jadwal perkuliahan yang padat dan kegiatan kemahasiswaan yang diambil oleh para mahasiswa menyebabkan mereka memiliki sedikit waktu untuk beristirahat. Dengan jadwal yang padat ini seiring aktivitas yang melelahkan tubuh kita memerlukan air yang cukup agar tubuh kita tidak mengalami dehidrasi (kekurangan air) (Sembiring, 2010).

**Uji Kenormalan Multivariat**

Uji kenormalan pada multivariat sebenarnya sangat komplek, karena harus dilakukan pada seluruh variabel secara bersama-sama. Namun uji ini bisa juga dilakukan pada setiap variabel, dengan logika bahwa jika secara individual masing-masing variabel memenuhi asumsi normalitas, maka secara bersama-sama (multivariat) variabel-variabel tersebut juga bisa dianggap memenuhi asumsi kenormalan dengan melihat besaran kolmogorov-smirnov atau melihat grafik dengan kriteria pengujian (Santoso, 2002).

Angka P-Value > 0,05 maka data berdistribusi normal.

Angka P-Value < 0,05 maka data tidak berdistribusi normal.

Berdasarkan teori Wahl & Kronmal (1977), seringkali kenormalan ganda sulit diperoleh terutama bila sampel yang diambil relatif kecil. Bila hal ini terjadi, uji vektor rata-rata tetap bisa dilakukan selama asumsi kedua (kesamaan kovariansi) dipenuhi, yaitu bila nilai uji kovarian data dengan melihat matrik kovariansi melalui tampilan nilai *Log Determinants* tidak berbeda jauh sehingga dapat diasumsikan bahwa matriks kovarian relatif sama untuk kedua kelompok (Mattjik dan Sumertajaya, 2011).

**Matrik Kovarian**

Misalkan *X* = (*X1, . . . . , Xp*)T adalah suatu vektor *p* dari variabel acak, disebut juga vektor acak. Masing-masing komponen *X*  adalah suatu variabel acak univariat *X1* ( *j* = 1, . . . , *p*) dengan distribusi mardinalnya mempunyai nilai μ*j* = *E* [ *Xj* ] dan varians $σ\_{j}^{2}$ = *E* $\left[( X\_{j}- μ\_{j}) \right]$. Nilai harapan *X* adalah sebagai vektor dari nilai-nilai yang diharapkan dari komponennya, yaitu:

 *E* [*X*] = $\left[\begin{matrix}E(X\_{1})\\E(X\_{2})\\\begin{matrix}\vdots \\E(X\_{p})\end{matrix}\end{matrix}\right]$ = $\left[\begin{matrix}μ\_{1}\\μ\_{2}\\\begin{matrix}\vdots \\μ\_{p}\end{matrix}\end{matrix}\right]$ = μ (1)

Matrik kovarian populasi *X*  digambarkan sebagai berikut :

 *Cov* $\left[X\right]$ = E $\left[\left(X\_{j}- μ\_{j}\right)(X\_{k}- μ\_{k})\right]$

 $\left\{\left[\begin{matrix}X\_{1}-μ\_{1} \\X\_{2}- π\_{2}\\\begin{matrix}\vdots \\X\_{P}- μ\_{P}\end{matrix}\end{matrix}\right] \left[X\_{1}- μ\_{1} X\_{2}- μ\_{2} \cdots X\_{P}-μ\_{P}\right]\right\}$=E $\left\{\left[\begin{matrix}X\_{1}-μ\_{1} \\X\_{2}- π\_{2}\\\begin{matrix}\vdots \\X\_{P}- μ\_{P}\end{matrix}\end{matrix}\right] \left[X\_{1}- μ\_{1} X\_{2}- μ\_{2} \cdots X\_{P}-μ\_{P}\right]\right\}$

*= E*$\left[\begin{matrix}(X\_{1}- μ\_{1})^{2}&\left(X\_{1}- μ\_{1}\right)(X\_{2}- μ\_{2})&…\\\left(X\_{2}- μ\_{2}\right)(X\_{1}- μ\_{1})&(X\_{2}- μ\_{2})^{2}&…\\\begin{matrix}\vdots \\\left(X\_{P}- μ\_{P}\right)(X\_{1}- μ\_{1})\end{matrix}&\begin{matrix}\vdots \\\left(X\_{P}- μ\_{P}\right)(X\_{2}- μ\_{2})\end{matrix}&\begin{matrix}\\…\end{matrix}\end{matrix} \begin{matrix}\left(X\_{1}- μ\_{1}\right)(X\_{P}- μ\_{P})\\\left(X\_{2}- μ\_{2}\right)(X\_{P}- μ\_{P})\\\begin{matrix}\vdots \\(X\_{P}-μ\_{P})^{2}\end{matrix}\end{matrix}\right]$ (2)

Untuk *j* = 1, 2 . . . ,*p* dan *k* = 1, 2, . . . ,*p* (Johnson dan Wichern, 2002).

Unsur-unsur diagonal sama dengan *E*$\left[(X\_{j}- μ\_{j})^{2}\right]$ = $σ\_{j}^{2}$. Unsur-unsur *off-diagonal* sama dengan *E*$\left[\left(X\_{j}- μ\_{j}\right)(X\_{k}-μ\_{k})\right]$ = *Cov* ( $X\_{j},X\_{k}$) = σjk adalah kovarian atau peragam antara variabel *Xj* dan *Xk*. Kemudian matrik kovarian dapat ditulis sebagai berikut (Johnson dan Wichern, 2002):

 *Cov* (*X*) = $\left[\begin{matrix}σ\_{11}&σ\_{12}&…\\σ\_{21}&σ\_{22}&…\\\begin{matrix}\vdots \\σ\_{p1}\end{matrix}&\begin{matrix}\vdots \\σ\_{p2}\end{matrix}&\begin{matrix}\ddots \\\end{matrix}\end{matrix} \begin{matrix}σ\_{p1}\\σ\_{p2}\\\begin{matrix}\vdots \\σ\_{pp}\end{matrix}\end{matrix}\right]$ (3)

**Uji Vektor Rata-Rata**

Statistika *V*-Bartlett dapat dipergunakan untuk menguji vektor-vektor nilai rata-rata di antara kelompok apakah berbeda atau tidak berbeda. Statistik *V* itu akan berdistribusi mendekati distribusi *chi-square* (χ2) dengan derajat bebas *p* (*h*-1), dimana *p* adalah banyaknya variabel pembeda dan *h* adalah banyak kelompok (populasi) pengamatan. Dengan demikian besar *V* dapat dibandingkan dengan tabel *chi-square* pada taraf *a* tertentu untuk menguji hipotesis tentang kesamaan vektor nilai rata-rata diantara populasi yang dipelajari. Aturan pengambilan keputusan adalah :

Jika *V* $\leq χ\_{a,v=p(h-1)}^{2}$, maka *H0* diterima.

Jika *V* $> χ\_{a,v=p(h-1)}^{2}$, maka *H0* ditolak

Penerimaan *H*0 berarti tidak ada perbedaan nilai rata-rata antara kelompok (populasi), sehingga dalam hal ini fungsi diskriminan tidak perlu dibentuk (tidak layak bentuk), sebaliknya penolakan *H*­0 berarti terdapat perbedaan nilai rata-rata diantara kelompok sehingga layak untuk mengkaji perbedaan dan dibangun fungsi diskriminannya. Jika suatu fungsi diskriminan memang layak dibentuk dan telah berhasil mendapatkan fungsi diskriminan itu.

**Model Analisis Diskriminan**

Pada dasarnya analisis diskriminan dapat dilakukan dengan menggunakan pendekatan matrik peragam dalam kelompok (*W*) dan matrik peragam antar kelompok (*B*). Fungsi diskriminan dibangkitkan sedemikian rupa sehingga dengan menggunakan skor diskriminan maka keragaman relatif antar kelompok terhadap keragaman dalam kelompok menjadi maksimum (Tatham, et.al, 1998).

Sebagai metode untuk membedakan objek menjadi dua atau lebih grup maka analisis diskriminan dapat dihitung dengan memberikan timbangan kepada setiap variabel independen agar bisa memaksimalkan perbedaan antara grup yang ada. Model analisis dikriminan berkenaan dengan kombinasi linier memiliki bentuk sebagai berikut (Supranto, 2004) :

*Y* = *b0* + *bi1 Xi1 + b2x2 + . . . + bp xp*  (4)

dengan :

*Y* = nilai (skor) diskriminan dan merupakan variabel terikat.

*Xj* = variabel (atribut) ke –*j* dan merupakan variabel bebas

*bj* = koefisien diskriminan/bobot dari variabel (atribut) ke -*j.*

Dikaitkan dengan persamaan (4), jika terdapat sampel acak berukuran *n* dari populasi *p* untuk *n* = 1, 2, ..., *j*. Kemudian diketahui *p* buah variabel *X*1*, X*2, ..., *Xp*, maka akan diperoleh matrik data pengamatan berukuran *p* x *nk* dari populasi *pk* (Johnson dan Wichern, 2002):

 $\begin{matrix}X\\(pxn\_{n})\end{matrix}$ =$ \left[\begin{matrix}x\_{1n1}&x\_{1n\_{2}}&…\\x\_{2n\_{2}}&x\_{2n\_{2}}&…\\\begin{matrix}\vdots \\x\_{pn\_{1}}\end{matrix}&\begin{matrix}\vdots \\x\_{pn\_{2}}\end{matrix}&\begin{matrix}\\…\end{matrix}\end{matrix} \begin{matrix}x\_{1n\_{k}}\\x\_{2n\_{k}}\\\begin{matrix}\vdots \\x\_{pn\_{k}}\end{matrix}\end{matrix}\right]$ (5)

Karena fungsi diskriminan didefinisikan sebagai *Y =aTX* seperti Persamaan (4), maka keragamaan antar kelompok dari *Y*  adalah *aTBa* dan keragaman dalam kelompok *Y* adalah *aTWa*. Selanjutnya λ didefinisikan sebagai keragaman relatif antara kelompok terhadap keragaman dalam kelompok yang didefinisikan (Gaspersz, 1995):

λ = $\frac{a^{T}Ba}{a^{T}Wa}$ $<=>$ λ *aT Wa = aT Ba* (6)

Berdasarkan persamaan (6), maka masalah dalam analisis diskriminan adalah bagaimana menentukan vektor pembobot *a* yang memaksimumkan λ dalam persamaan (6). Berdasarkan aturan kalkulus, untuk memperoleh vektor pembobot *a* yang memaksimumkan λ, maka dapat diturunkan :

 $\frac{∂λ}{∂a}$ = $\frac{\left\{\left(2Ba\right)\left(a^{T}Wa\right)-\left(2Wa\right)(a^{T}Ba)\right\}}{(a^{T}Wa)^{2}}$

 = $\frac{2\left\{\left(Ba\right)\left(a^{T}Wa\right)-\left(Wa\right)(a^{T}Ba)\right\}}{(a^{T}Wa)^{2}}$

 =$\frac{2\left\{\left(Ba\right)\left(a^{T}Wa\right)-\left(Wa\right)(λa^{T}Wa)\right\}}{(a^{T}Wa)^{2}}$

 =$\frac{2(Ba- λWa)}{(a^{T}Wa)}$ = 0

Hasil turunan di atas akan sama saja dengan persamaan berikut (Johnson dan Wichern, 2002):

 (*B – λW*) *a=0* (7)

Dengan demikian vektor koefisiensi akan diperoleh melalui penyelesaian persamaan (7).

Matrik *B* berpangkat (mempunyai rank) min(*h-*1*,p*), tergantung mana yang terkecil apabila banyaknya kelompok dikurang 1 atau banyaknya variabel pembeda, sedangkan matrik *W* berpangkat penuh (*full rank*) yaitu sebesar *p* (banyaknya variabel pembeda yang dipelajari). Dengan mengasumsikan matrik *W* adalah *nonsingular* sehingga dapat ditentukan inversnya, yaitu *W-*1*,* maka diperoleh persamaan (7), digandakan awalnya dengan *W*-1 akan diperoleh (Gaspersz, 1995):

(*W*-1 *B- λI*)*a =0* (8)

Pada dasarnya koefesien pembobot fungsi diskriminan *aT* diperoleh bedasarkan persamaan (Gaspersz, 1995):

(*B –* λ*W*)a *=* (*W-1 B – λI*)*a =0*  (9)

dengan :

*B* = matrik kovarian antara kelompok

*W* = matrik kovarian dalam kelompok

λ = akar ciri (*eigenvalue*)

*a =* vektor ciri

Persamaan (9) baru akan mempunyai solusi apabila dipenuhi syarat determinan |*W-1 B – λI*| = 0. Jika Persamaan (9) diberikan batasan *aT Wa* = 1, maka akan ditemukan solusi dan membuat *λ = aTBa*. Dengan demikian tampak bahwa nilai maksimum *λ* akan menerangkan keragaman antar kelompok yang maksimum. Selanjutnya apabila diberikan kendala lain, yaitu : $a\_{1}^{T}$*Ba2* = 0 maka menurut Tatham, et.al (1998), akan membuat fungsi diskriminan pertama tidak berkorelasi dengan fungsi diskriminan kedua.

**Pembentuk Fungsi Diskriminan**

Dalam analisis diskriminan untuk menghilangkan pengaruh perbedaan satuan pengukuran terhadap besaran koefisien pembobot fungsi diskriminan, maka biasanya ditentukan pula koefisien bakunya, menggunakan persamaan berikut (Gaspersz, 1995):

$$a\_{mi}^{\*}= \sqrt{w\_{ii}} a\_{m} :i=1,2,…,P$$

 *m* = 1,2, $…$ ,*s*

 *s* = min (*h -*1,*p*) (10)

 dengan :

$a\_{mi}^{\*}=$koefisien diskriminan baku dari variabel ke-*i* dalam fungsi diskriminan ke-*m*.

*wii* = elemen diagonal utama dalam matrik *W*

*ami =* koefisien diskriminan tak baku dari ke-*i* dalam fungsi diskriminan ke- *m.*

Untuk mengetahui beberapa banyak fungsi diskriminan yang berkontribusi nyata dalam menerangkan perbedaan diantara kelompok yang ada, maka diperlukan suatu uji statistik. Proses pengujian dilakukan dengan jalan mengkaji apakah diskriminan sisa (*residual discrimination*) setelah diterangkan oleh fungsi diskriminan pertama, fungsi diskriminan pertama dan kedua, dan seterusnya bersifat nyata dalam statistik. Jika hasil pengujian bersifat nyata berarti masih perlu dibentuk fungsi diskriminan berikutnya, sedangkan apabila pengujian diskriminasi sisa tidak nyata berarti perlu dibentuk fungsi diskriminan berikutnya, pengujian statistik secara berurutan ini serta derajat bebasnya ditunjukan dalam Tabel 1 (Gaspersz, 1995).

Tabel 1. Pengujian Diskriminasi Sisa Dalam Analisis Diskriminan

|  |  |  |
| --- | --- | --- |
| Diskriminasi sisa | Perkiraan statistikχ2 | Derajat bebas |
| Fungsi Diskriminan 1 | *V – V1* | *p*(*h –* 1) *–* (*p+h–*2)*=*(*p–*1)(*h–*2) |
| $$\vdots $$ | $$\vdots $$ | $$\vdots $$ |

dengan :

*p* : banyaknya variabel dalam fungsi diskriminan

*h* : banyaknya kelompok (populasi)

**Keeratan Hubungan Fungsi Diskriminan**

Untuk mengetahui keeratan hubungan fungsi diskriminan dapat diketahui dengan melihat nilai koefisien korelasi kanonikal yang merupakan ukuran keeratan hubungan fungsi diskriminan dengan (*h -*1) variabel *dummy* yang membatasi anggota-anggota *h* kelompok, dihitung berdasarkan persamaan :

*Rcm* = $\sqrt{λ\_{m}/(1+ λ\_{m} ) }$ (11)

dengan:

*Rcm* = korelasi kanonik antara fungsi diskriminan ke-*m* dan (*h-*1) variabel *dummy* yang membatasi anggota-anggota dari *h* kelompok.

λm = *eigenvalue* ke-*m*, *m* = 1,2...,s,

s = min(*h-*1*,p*), *h* = banyaknya kelompok dan *p* = banyaknya variabel pembeda dalam fungsi diskriminan.

**Peranan Relatif Fungsi Diskriminan**

Peranan relatif suatu fungsi diskriminan ke-*i* dalam memisahkan anggota-anggota kelompok diukur dari persentasi relatif *eigenvalue* yang berhubungan dengan fungsi diskriminan tersebut sehingga dapat ditentukan sebagai berikut (Gaspersz, 1995):

Peranan relatif *Yi =* $\frac{λ\_{i}}{\sum\_{m=1}^{S}λ\_{m}} x 100\%$ (12)

*s* = min (*h*-1,p)

 Jumlah *eigenvalue* yang berhubungan dengan fungsi diskriminan merupakan ukuran keseragaman total dari variabel diskriminan.

**Analisis Diksriminan Metode Linier Fisher**

Prinsip utama dari fungsi diskriminanlinier Fisher adalah pemisahan sebuah populasi. Fungsi diskriminan yang terbentuk dapat digunakan untuk pengelompokkan suatu observasi berdasarkan kelompok-kelompok tertentu. Metode Fisher ini mengasumsikan data harus berdistribusi normal, tapi dalam perhitungan salah satu syarat yang harus diperhatikan adalah data yangdigunakan harus memiliki matrik kovarian yang sama untuk setiap kelompok populasi yang diberikan. Kriteria penggolongan berdasarkan persamaan diskriminan linier Fisher, ditunjukan pada Persamaan menurut (Suryanto, 1990). Alokasikan matrik kovarian kelompok (populasi) *k* (*k =* 1, 2, *. . . ,h*) jika

$\sum\_{m=1}^{r}(y\_{m}- \overbar{y}\_{km})^{2} $=$\sum\_{m=1}^{r}[ a\_{m}^{T}(x- \overbar{x}\_{k})]^{2} \leq \sum\_{m=1}^{r}[ a\_{m}^{T}(x- \overbar{x}\_{h})]^{2}$ (13)

Untuk semua *h ≠ k* ; *r* ≤ *s; s = min* (*h-1,p*) dengan:

*ym* : skor diskriminan ke-*m* dari objek

$\overbar{y}\_{m}$: rata-rata skor diskriminan ke –*m* dari kelompok ke –*k*  (*k* = 1, 2, ..., *h*).

$a\_{m}^{T} $: vektor koefisien fungsi diskriminan ke –*m*

*x*  : vektor data pengamatan dari objek yang akan digolongkan.

$\overbar{x}\_{k} $ : vektor nilai rata-rata variabel pembeda dari kelompok ke-*k.*

 *r*  : banyaknya fungsi diskriminan yang

digunakan dalam penggolongan.

*s* : banyaknya fungsi diskriminan yang mugkin dibentuk dalam analisis diskriminan.

 *h* : banyaknya kelompok (populasi) yang dipelajari.

*p* : banyaknya variabel pembeda dalam analisis diskriminan.

***Cut off Score***

Proses ini dilakukan untuk membuat nilai batas untuk mengetahui konsumen masuk ke dalam grup sedikit mengkonsumsi AMDK atau banyak mengkonsumsi AMDK dengan menggunakan rumus sebagai berikut (Santoso, 2002):

Zcu =$\frac{N\_{A}Z\_{A}+ N\_{B}Z\_{B}}{n}$ (14)

dengan :

Zcu  = angka kritis yang berfungsi sebagai nilai batas

 NA = jumlah sampel grup A

 NB = jumlah sampel grup B

 ZA = nilai *centroid* grup A

 ZB = nilai *centroid* grup B

 *n* = banyak data

**Penilaian Validitas Diskriminan**

Bobot diskriminan diperkirakan dengan menggunakan *analysis sampel* dikalikan dengan nilai variabel bebas di dalam *holdout sampel* untuk mendapatkan skor diskriminan untuk kasus yang berada dalam *holdout sampel.* Objek atau kasus tersebut kemudian dimasukan ke dalam kelompok berdasarkan pada nilai fungsi diskriminan dan aturan-aturan yang tepat (Saparita, 1997). Untuk lebih jelasnya dapat dilihat pada Tabel 2 di bawah ini.

Tabel 2. Klasifikasi *Actual Group* (kelompok sebenarnya) dan *Predicted Group* (Kelompok yang diprediksi)

|  |  |  |
| --- | --- | --- |
|  |  | *Predicted Group* |
|  |  | (Kelompok yang diprediksi) |
|  |  | π1 |  π2 |
| *Actual Group*(kelompok sebenarnya) | *π*1 | *n*1*c* | *n*1*M = n*1 *– n*1*C* |
| *π*2 | *n*2*M = n*2*– n*2*C* | *n*2*C* |

Persentase ketepatan klasifikasi = $\frac{n\_{1c}+ n\_{2c}}{n} ×100\% $ dimana *n = n*1*c+ n*2*c+ n*1*m+n*2*m*

dengan :

*n*1*C*= banyak pengamatan π1 yang

dikelompokkan secara benar sebagai π1

*n*1*M**=* banyak pengamatan π1 yang salah

dikelompokkan sebagai π2

*n*2*C* = banyak pengamatan π2 yang

dikelompokkan secara benar sebagai π2

*n*2*M* = banyak pengamatan π2 yang salah

dikelompokkan sebagai π

Menurut Santoso (2002). jika nilai ketepatan di atas 50%, ketepatan model dianggap tinggi.

**Metodologi Penelitan**

Data yang digunakan dalam penelitian ini adalah data kuesioner Mahasiswa/i Fakultas Matematika Dan Ilmu Pengetahuan Alam Universitas Mulawarman Program Studi Statistika angkatan 2012. Data tersebut kemudian dipergunakan sebagai variabel dalam melakukan analisis diskriminan. Variabel dalam penelitian ini terdiri dari variabel dependen dan independen sebagai berikut:

1. **Variabel Dependen**

Variabel dependen yang digunakan dalam penelitian ini ada dua kategori, yaitu:

1. Sedikit (konsumen yang termasuk tipe sedikit minum air mineral yang dimaksud ke dalam sedikit bila dalam 1 (satu) bulan mengkonsumsi kurang dari empat kali).
2. Banyak (konsumen yang termasuk tipe banyak minum air mineral yang dimaksud ke dalam banyak adalah apabila dalam 1 (satu) bulan mengkonsumsi lebih dari atau sama dengan empat ($\geq $4) kali).
3. **Variabel Independen**

Penulis menggunakan beberapa variabel dalam penelitian ini yaitu:

1. Usia konsumen (tahun).
2. Berat badan konsumen (kilogram).
3. Tinggi badan konsumen (centimeter).
4. Uang Saku (Rupiah/bulan).
5. Berapa jam kegiatan di kampus dalam sehari (jam).
6. Kegiatan olahraga konsumen dalam seminggu (jam).

**Hasil dan Pembahasan**

**Uji Kenormalan Multivariat**

Untuk mengetahui apakah variabel berdistribusi normal dapat diketahui dengan menlihat nilai P-Value < 0,05 atau sebaliknya dengan perumusan hipotesis sebagai berikut :

H0 : variabel-variabel yang diteliti berdistribusi normal

H1 : variabel-variabel yang diteliti tidak berdistribusi normal

Perhitungan dengan menggunakan bantuan program SPSS diperoleh hasil seperti pada Tabel 3.

Berdasarkan Tabel 3, terlihat nilai P-Value untuk setiap variabel nilainya lebih kecil dari 0,05 (P-Value < 0,05), kecuali pada variabel tinggi yaitu P-Value > 0,05. Sehingga h0 ditolak yang artinya data tidak berdistribusi normal, namun seringkali kenormalan ganda sulit diperoleh terutama bila sampel yang diambil relatif kecil. Bila hal ini terjadi, uji vektor rata-rata tetap bisa dilakukan selama asumsi kedua (kesamaan kovarian) dipenuhi, yaitu bila nilai uji kovarian data dengan melihat matrik kovariansi melalui tampilan nilai *Log Determinants* tidak berbeda jauh atau dapat diasumsikan bahwa matriks kovariansi relatif sama untuk kedua kelompok.

 Tabel 3. Uji Kenormalan Multivariat

|  |  |
| --- | --- |
| Variabel | Kolmogorov-Smirnov |
| Statistik | df | P-Value |
| Minum | 0,357 | 60 | 0,000 |
| Usia | 0,359 | 60 | 0,000 |
| Berat | 0,216 | 60 | 0,000 |
| Tinggi | 0,111 | 60 | 0,064 |
| Uang Saku | 0,151 | 60 | 0,002 |
| Di Kampus | 0,237 | 60 | 0,000 |
| Olahraga | 0,142 | 60 | 0,000 |

**Uji Matriks Kovarian**

Pengujian ini dilakukan dengan melihat melalui tampilan nilai *Log Determinants* apakah antar kelompok nilainya sama atau tidak. Kemudian diperoleh Tabel 4.

Tabel 4. *Log Determinants*

|  |  |  |
| --- | --- | --- |
| Variabel Minum | Peringkat | *Log Determinants* |
| Sedikit | 6 | 18,106 |
| Banyak | 6 | 18,253 |
| Pemusatan dalam kelompok | 6 | 18,663 |

Berdasarkan Tabel 4, angka log *determinants* untuk kategori sedikit adalah 18,106 dan kategori banyak 18,253 yang artinya *Log Determinants* tidak berbeda jauh sehingga dapat diasumsikan bahwa matrik kovarian relatif sama untuk kedua kelompok.

**Uji Vektor Rata-Rata**

Pengujian kesamaan vektor rata-rata dari kedua variabel untuk mengetahui apakah variabel memiliki perbedaan antar kelompok. Pengujian dapat dilakukan menggunakan uji V-Bartlett dengan hipotesis sebagai berikut;

H0 : variabel-variabel yang diteliti tidak memiliki perbedaan antar kelompok.

H1 : variabel-variabel yang diteliti memiliki

perbedaan antar kelompok.

Kriteria pengujian H0 ditolak jika nilai χ2hitung  > χ2tabel atau nilai P-Value < 0,05.

Tabel 5. Besaran Nilai Khi Kuadrat

|  |  |  |
| --- | --- | --- |
| Diskriminasi | Khi Kuadrat | Db |
| V |  17,764 |  4 |

Karena uji V-Bartlett berdistribusi mendekati distribusi Khi Kuadrat, dari Tabel 5 diperoleh hasil nilai khi Kuadrat sebesar 17,764, dengan derajat bebas *v=p*(*g-*1)=6(2-1)*=*6. Mengambil taraf nyata α = 5% dari Tabel distribusi Khi Kuadrat sebesar 12,592. Karena 17,764 > 12,592 atau nilai P-Value < 0,05 , maka Ho ditolak sehingga dapat disimpulkan bahwa variabel-variabel yang diteliti memiliki perbedaan antar kelompok.

**Fungsi Diskriminan**

Menjelaskan hubungan antara variabel bebas dengan fungsi diskriminan yang terbentuk pada proses perhitungan dilakukan dengan bantuan program SPSS 20.0 diperoleh hasil pada Tabel 6.

Tabel 6. Keeratan Hubungan Variabel Bebas Dengan

 Fungsi Diskriminan

|  |  |
| --- | --- |
|  | Fungsi |
| 1 |
| Di kampus | 0,473 |
| Olahraga | 0,459 |
| Usia | -0,435 |
| Uang saku | 0,433 |
| Berata | 0,062 |
| Tinggia | -0,048 |

Dari Tabel 6 menjelaskan korelasi antara variabel bebas dengan fungsi diskriminan yang terbentuk. Terlihat bahwa variabel kegiatan di kampus dalam sehari paling erat hubungannya dengan fungsi diskriminan 0,473 diikuti oleh variabel olahraga, usia, uang saku. Hanya saja, untuk variabel berat dan tinggi tidak dimasukan dalam model diskriminan.

 Kemudian untuk mengetahui fungsi diskriminan yang terbentuk dapat dilihat pada Tabel 7.

Tabel 7. Fungsi Diskriminan

|  |  |
| --- | --- |
| Variabel | Fungsi |
| 1 |
| Usia | -0,752 |
| Uang saku | 0,004 |
| Di Kampus | 0,558 |
| Olahraga | 0,301 |
| (Nilai konstan) | 7,557 |

Kemudian diperoleh fungsi diskriminan sebagai berikut :

*Y* = 7,557 - 0,752 usia + 0,004 uang saku + 0,558 di kampus + 0,301 olahraga. (14)

**Keeratan Hubungan Fungsi Diskriminan**

Proses ini dilakukan untuk mengetahui korelasi kanonik antara fungsi diskriminan kemudian diperoleh :

Tabel 8. Korelasi Kanonik

|  |  |
| --- | --- |
| Fungsi | Korelasi Kanonik |
| 1 | 0,521 |

Untuk Tabel 8, korelasi kanonikmengukur keeratan hubungan antara skor diskriminan dengan kelompok (dalam hal ini, karena ada dua tipe konsumen maka ada dua kelompok). Angka 0,521 menunjukan keeratan yang cukup bagus, dengan ukuran skala 0 sampai 1.

**Peranan Relatif Fungsi Diskriminan**

Dari persentase relatif akar ciri (*eigenvalue*) yang berhubungan dengan fungsi diskriminan, kemudian dilakukan perhitungan dengan menggunakan bantuan program SPSS versi 20.0 diperoleh hasil pada Tabel 9.

Tabel 9. Peranan Relatif

|  |  |
| --- | --- |
| Fungsi | Peranan Relatif % |
| 1 | 100,0 |

Terlihat bahwa persentase relatif yang dapat dijelaskan oleh fungsi diskriminan adalah sebesar 100,0 %. Yang berarti proses pemisahan anggota kelompok terjadi sempurna dengan persentase sebesar 100%.

**Pengklasifikasian**

Pengklasifikasian berdasarkan diskriminan linier Fisher adalah untuk membandingkan skor diskriminan dari suatu objek pengamatan terhadap skor rata-rata kelompok dan suatu objek pengamatan digolongkan ke dalam kelompok tertentu.

Tabel 10. Fungsi Diskriminan Berdasarkan Kelompok

|  |  |
| --- | --- |
|  | Minum |
| sedikit | banyak |
| Usia | 38,851 | 37,946 |
| Uang saku | 0,002 | 0,006 |
| Di kampus | 13,365 | 14,037 |
| Olahraga | -1,258 | -0,895 |
| (Nilai tetap) | -387,373 | -378,322 |

Dari Tabel 7 fungsi diskriminan dari Fisher sama dengan fungsi diskriminan sebelumnya, hanya saja pembagian berdasarkan kode kelompok.

Kemudian dicari selisih antara sedikit dan banyak, untuk melihat bahwa tidak berbeda jauh antara fugsi diskriminan yang terbentuk dengan fungsi diskriminan Fisher, kemudian diperoleh hasil yant tidak berbeda jauh antara nilai fungsi diskriminan persamaan (14) dengan hasil perhitungan selisih antara fungsi diskriminan Fisher dibawah ini:

 *Y* = 9,051 + 0,905 usia + 0,0004 uang saku + 0.672 di kampus + 0,0363 (15)

Dengan membandingkan fungsi diskriminan dapat dilihat bahwa bentuk persamaan tersebut hampir sama dengan Tabel 10, akan tetapi yang digunakan adalah skor fungsi diskriminan Tabel 7 bukan selisih antara kategori. Fungsi pada kelompok centroidnya dapat dilihat pada Tabel 11.

Tabel 11. Fungsi Pada Kelompok Centroid

|  |  |
| --- | --- |
| Minum | Fungsi |
| 1 |
| Sedikit | -0,562 |
| Banyak | 0,642 |

Oleh karena ada dua tipe konsumen, maka disebut *Two-Group Discriminant.* Dimana kelompok yang satu mempunyai *centroid* (*group means*) negatif, dan kelompok lainnya mempunyai *centroid* (*Group means*) positif. Dapat dilihat pada Tabel 12.

Tabel 12. Grup Centroid

|  |  |  |  |
| --- | --- | --- | --- |
| Minum | Persentase % | Sampel | Grup Centroid |
| Sedikit | 50 | 32 | -0,562 |
| Banyak | 50 | 28 | 0,642 |

Tabel 12 dapat juga digunakan untuk perhitungan *cutting score* (nilai batas) sebagai berikut :

 Zcu =$\frac{N\_{A}Z\_{A}+ N\_{B}Z\_{B}}{n}$ = $\frac{(32 x -0,562) +(28 x 0,642) }{60}$

 =$ \frac{-17,984 +17,976}{60}$ = 0,008 (16)

 Penggunaan angka ZCU­ (diskriminan Z skor) : jika skor nilainya untuk konsumen sedikit pada Tabel 12 di atas ZCU, maka masuk ke dalam kelompok banyak. Sebaliknya jika nilai di bawah ZCU maka masuk ke dalam kelompok sedikit.

**Penilaian Validitas Diskriminan**

Selanjutnya akan dilihat seberapa besar hasil klasifikasi tersebut tepat. Dengan kata lain berapa persen kesalahan klasifikasi pada proses klasifikasi tersebut. Dengan bantuan program SPSS diperoleh hasil seperti pada Tabel 13.

Dari Tabel 13 terlihat bahwa konsumen yang pada awal tergolong ke dalam kelompok sedikit dan dari klasifikasi model diskriminan tetap pada kelompok sedikit adalah 26 orang, sedangkan konsumen yang awalnya masuk ke dalam kelompok sedikit ternyata menjadi kelompok banyak berjumlah 6 orang. Demikian juga dengan kelompok banyak, yang tetap masuk kelompok banyak berjumlah 20 orang dan yang masuk menjadi kelompok sedikit adalah 8 orang.

Tabel 13. Validitas Diskriminan

|  |  |  |
| --- | --- | --- |
| Minum | Prediksi Anggota Grup | Total |
| Sedikit | Banyak |
| Data sebenarnya | Sedikit | 26 | 6 | 32 |
| Banyak | 8 | 20 | 28 |
| Persentasi | Sedikit | 81,3 | 18,8 | 100 |
| Banyak | 32,1 | 71,4 | 100 |

Ketepatan model dapat diketahui dengan perhitungan persentasi ketepatan klasifikasi :

Ketepatan klasifikasi = $\frac{n\_{1c}+ n\_{2c}}{n} $x 100%

 =$ \frac{26+20}{60}$ x 100% = 76 (17)

Karena angka ketepataan prediksi dari model diskriminan sebesar 76% > 50%. Menurut Santoso (2002), jika nilai ketepatan di atas 50%, ketepatan model dianggap tinggi.

**Kesimpulan**

Berdasarkan hasil pengolahan data konsumen air mineral dalam kemasaan (AMDK) dapat ditarik kesimpulan penelitian sebagai berikut:

Model fungsi diskriminan yang terbentuk berdasarkan hasil pengolahan data adalah :

*Y* = 7,557 - 0,752 usia + 0,004 uang saku + 0,558 kegiatan di kampus + 0,301 kegiatan olahraga. (18)

Dari 60 data atau 100% data yang digunakan dimana 32 konsumen masuk ke dalam kategori sedikit mengkonsumsi (AMDK) dan 28 konsumen masuk ke dalam kategori banyak mengkonsumsi (AMDK). Kemudian dilakukan proses ketepatan pengklasifikasi pada persaman (17), diperoleh hasil pengklasifikasian sebesar 76%, dimana ternyata hanya 26 konsumen yang masuk ke dalam kategori sedikit mengkonsumsi (AMDK) dan 20 konsumen masuk dalam kategori banyak mengkonsumsi (AMDK).

**Daftar Pustaka**

Gaspersz, Vincent. 1995. Tehnik Analisis Dalam Penelitian Percobaan. Tarsito: Bandung.

Johnson, Richard A.,Dean W. Wichern. (2002). *Applied Multivariate Statistical Analysis (Third Edition)*. New Jersey.

Mattjik, Ansori Ahmad., Sumertajaya, I Made. (2011). *Sidak Peubah Ganda Dengan Menggunakan SAS*. Bogor. IPB PRESS. [Online].

Santoso, Singgih. (2002). *Buku Latihan SPSS Staistik Multivariat.* Jakarta: PT. Elex Media komputindo.

Saparita, Rachmini. (1997). Analisis Diskriminan Sebagai Penentu Peubah Ciri Kelompok Populasi. [Online].

Sembiring, Amstrong. (2010). *Menyoal masyarakat konsumen air* . Dikutip dari <<http://sosbud.kompasiana.com/2010/0124/menyoal-masyarakat-konsumen-air/>>. Tanggal akses 2 Desember 2012.

Supranto, J. (2004). *Analisis Multivariat Arti Dan Interprestasi.* Jakarta: PT. Rieneka Cipta.

Suryanto. (1990). *Metode Statistika Multivariat*. Jakarta:P2LPTK.

Tatham, Et, Al. (1998). *Multivariate Data* *Analysis*. Prentice Hall, New Jersey. [Online].