BAB II

TEORI PENUNJANG

2.1 VARIABEL DAN DATA

2.1.1 Jenis Variabel

Variabel merupakan sebuah konsep yang mempunyai bermacam-macam nilai. Misal dari konsep badan dapat dibuat bermacam-macam variabel, seperti berat badan, tinggi badan dan besar badan. Pada umumnya, variabel dibagi menjadi dua jenis yaitu variabel diskrit dan variabel kontinu.

1. Variabel kontinu

Variabel kontinu adalah variabel yang hasil pengukurannya (kodomain) berupa bilangan riil. Contoh berat badan, maka dapat diukur berat badan seseorang dan dicatat hasilnya. Ada yang 35,25 kg, 70 kg atau 20,45 kg. Variabel kontinu dapat dibagi menjadi tiga variabel yaitu ordinal, interval dan rasio. Variabel nominal tidak termasuk, karena hasil pengukuran variabel nominal hanya menunjukkan lambang (label) yang digunakan untuk mengklasifikasikan suatu objek, orang atau sifat.

a. Variabel ordinal, yaitu variabel yang menunjukkan tingkatan-tingkatan. Misal panjang, kurang panjang, pendek. Dengan kata lain adalah variabel "lebih kurang" karena yang satu mempunyai kelebihan dibanding yang lainnya, sehingga dari tinjauan lain yang satu kurang dibanding yang lain.

Contoh : Ani terpandai, Rini pandai dan Siti tidak pandai.

b. Variabel interval, yaitu variabel yang dua satuan pengukuran yang berbeda mempunyai arti yang sama. Contoh pengukuran temperatur (suhu). Derajat temperatur didefinisikan dengan perubahan volume air raksa dalam termometer. Jika dua volume air raksa dalam termometer, maka menunjukkan temperatur dua derajat.

c. Variabel ratio, yaitu variabel perbandingan. Variabel ini dalam hubungan antar sesamanya merupakan "sekian kali".

   Contoh : Berat pak Ali 70 kg, sedangkan anaknya 35 kg. Maka pak Ali beratnya dua kali anaknya.

2. Variabel Diskrit

Variabel diskrit adalah variabel yang hasil pengukurannya (kodomain) berupa bilangan bulat. Variabel ini sering juga dinyatakan sebagai variabel kategori. Jika mempunyai dua kategori disebut variabel dikhotom sedangkan jika lebih dari dua kategori disebut variabel politom. Contoh variabel dikotom yaitu jenis kelamin, status perkawinan sedangkan contoh variabel politom adalah tingkat pendidikan.

2.1.2 DATA

Data merupakan hasil pencatatan sipeneliti, baik yang berupa fakta maupun angka. Menurut sifatnya data dibagi dua yaitu :

1. Data kualitatif, yaitu data yang tidak berbentuk angka. Misal sangat bagus, bagus, jelek dan sangat jelek. Dalam menganalisis data-data seperti ini diganti

dengan simbol angka. Misal sangat bagus = 1, bagus = 2, jelek = 3 dan sangat jelek = 4.

2. Data kuantitatif, yaitu data yang berbentuk angka. Data ini berasal dari data-data variabel diskrit dan kontinu.

2.2 VARIABEL RANDOM

Seandainya Ω menunjukkan ruang sampel, ξ menyatakan events dan P[.] merupakan fungsi peluang dengan domain ξ dan konterdomain interval [0,1], maka ruang peluang didefinisikan dengan (Ω,ξ,P[.]). Sehingga definisi variabel random adalah sebagai berikut.

Definisi 2.2.1 :

Diberikan ruang peluang (Ω,ξ,P[.]), variabel random dinotasikan dengan X atau X(.) adalah fungsi dengan domain Ω dan konterdomain garis riil. Fungsi X(.) sedemikian hingga himpunan A_r didefinisikan $A_r = \{w: X(w) \leq r\}$ berada dalam ξ untuk setiap bilangan r.

contoh 1 :

Pada percobaan pelemparan koin sekali. Variabel random X menyatakan jumlah gambar yang muncul. Ω = {gambar, angka} X(w) = 1 jika w = gambar dan X(w) = 0 jika w = angka. Akan ditunjukkan bahwa $\{w: X(w) \leq r\}$ berada dalam ξ untuk setiap bilangan riil r. ξ terdiri dari empat himpunan bagian yaitu : ϕ, {angka}, {gambar} dan Ω. Jika $r < 0$, $\{w : X(w) \leq r\} = \phi$, dan jika $0 \leq r < 1$, $\{w : X(w) \leq r\}$ = {angka} dan jika $r \geq 1$, $\{w : X(w) \leq r\} = \Omega$. Karena untuk setiap r himpunan $\{w : X(w) \leq r\}$ berada dalam ξ, maka X(.) adalah variabel random.

2.2.1 Variabel Random Diskrit

Definisi 2.2.2 :

Variabel random X didefinisikan diskrit jika range X adalah terhitung. Jika variabel random X diskrit, maka fungsi distribusi kumulatif F(x) adalah diskrit.

Definisi 2.2.3 :

Jika X variabel random diskrit dengan nilai $x_1, x_2, \ldots, x_n$, maka fungsi dinotasikan dengan f(.) dan didefinisikan dengan :

$$
\begin{aligned}
f(x) &= P[X = x] \quad && \text{jika } x = x_j \text{ , } j = 1, 2, \ldots, n, \ldots \\
&= 0 && \text{jika } x \neq x_j
\end{aligned}
$$

adalah didefinisikan untuk fungsi kepadatan diskrit X.

Nilai-nilai dari variabel random diskrit sering disebut titik massa, dan $f(x_j)$ menotasikan massa yang diasosiasikan dengan titik massa x_j. Fungsi massa peluang, fungsi frekuensi diskrit dan fungsi peluang adalah bentuk lain penggunaan fungsi kepadatan diskrit.

Definisi 2.2.4 :

Fungsi f(.) dengan domain garis riil dan konterdomain [0,1] didefinisikan untuk fungsi kepadatan diskrit jika untuk beberapa himpunan terhitung $x_1, x_2, \ldots, x_n, \ldots$

a. $f(x_j) > 0$ untuk $j = 1, 2, \ldots$

b. $f(x) = 0$ untuk $x \neq x_j$; $j = 1, 2, \ldots$

c. $\sum f(x_j) = 1$, dimana jumlahan untuk seluruh titik $x_1, x_2, \ldots, x_n, \ldots$

Definisi 2.2.5 :

Jika $X_1, X_2, \ldots, X_n$ variabel random dengan masing-masing fungsi peluang $f(x_1), f(x_2), \ldots, f(x_n)$ dan dengan fungsi peluang bersama $f(x_1, x_2, \ldots, x_n)$. Maka $X_1, X_2, \ldots X_n$ saling bebas jika $f(x_1, x_2, \ldots, x_n) = f(x_1). f(x_2)\ldots f(x_n)$ untuk semua kombinasi nilai $x_1, x_2, \ldots, x_n$.

Definisi 2.2.6 :

Jika X variabel random dengan fungsi peluang f(x), dan u(X) fungsi bernilai riil dari X, maka nilai ekspektasi u(X) ditulis dengan $E[u(X)]$ dengan :

$$E[u(X)] = \sum_{x} u(x)f(x)\text{, untuk seluruh kemungkinan nilai X} \qquad (2.2.1)$$

Jika jumlahan pada persamaan (2.2.1) infinite (tidak ada) maka dikatakan bahwa nilai ekspektasi dari u(X) tidak ada.

Dari definisi ini, akan dibicarakan dua macam nilai ekpektasi yaitu mean dan varian X.

Definisi 2.2.7 :

Jika X variabel random dengan fungsi peluang f(x), maka mean X dinotasikan dengan μ. Dimana $\mu = \sum_{x} xf(x)$ (2.2.2)

Definisi 2.2.8 :

Jika X variabel random dengan mean μ dan fungsi peluang f(x) varian X dinotasikan dengan σ^2 atau Var(X) adalah : $\sigma^2 = E[(X-\mu)^2] = E(X^2) - \mu^2$

Contoh 2 :

Misalkan sebuah kontraktor memerlukan X hari untuk menyelesaikan suatu pekerjaan yang ditawarkan, dimana X sebuah variabel random yang menyatakan jumlah hari untuk menyelesaikan pekerjaan. Distribusi peluang dari X yaitu (x,f(x)) adalah sebagai berikut :

x	f(x)
3	1/8
4	5/8
5	2/8

Dengan menggunakan rumus ekspektasi, dapat dihitung rata-rata dan varian dari X sebagai berikut :

$E(X) = \sum xf(x) = 3.1/8 + 4.5/8 + 5.2/8 = 33/8$

$Var(X) = E(X^2) - \mu^2$

$= \sum x^2 f(x) - \mu^2$

$= \left(3^2.1/8 + 4^2.5/8 + 5^2 2/8\right) - (33/8)^2 = 23/64$

Theorema 2.2.1 :

Jika $X_1, X_2, \ldots, X_n$ variabel random dan jika $Y = X_1+X_2+\ldots+X_n$ maka $E(Y) = E(X_1)+E(X_2)+\ldots+E(X_n)$.

Bukti :

$Y = X_1+X_2+\ldots+X_n$

$E(Y) = E(X_1+ X_2+ \ldots +X_n)$

$$= \sum\left(x_1 + x_2 + \ldots + x_n\right)f\left(x_1, x_2, \ldots x_n\right)$$
$$= \sum x_1 f\left(x_1, x_2, \ldots, x_n\right) + \ldots + \sum x_n f\left(x_1, x_2, \ldots, x_n\right)$$
$$= E(X_1) + \ldots + E(X_n)$$

Definisi 2.2.9 :

Jika X_1 dan X_2 variabel random dengan mean μ_1 dan μ_2, fungsi peluang $f(x_1)$ dan $f(x_2)$ dan fungsi peluang bersama $f(x_1,x_2)$. Covarian dari X_1 dan X_2 adalah :

$$\begin{aligned}\mathrm{Cov}(X_1,X_2) &= E[(X_1-\mu_1)(X_2-\mu_2)] \\ &= E[X_1X_2]-\mu_1\mu_2\end{aligned}$$

Theorema 2.2.2 :

Jika X_1 dan X_2 variabel random independent, maka covarian X_1 dan X_2 adalah nol.

Bukti :

$$\begin{aligned}\mathrm{Cov}(X_1,X_2) &= E(X_1,X_2) - \mu_1\mu_2 \\ &= \sum_{x1,x2} x_1x_2 f(x_1x_2)-\mu_1\mu_2 \\ &= \sum_{x1,x2} x_1x_2 f(x_1)f(x_2)-\mu_1\mu_2 \\ &= \sum_{x1} x_1 f(x_1)\sum_{x2} x_2 f(x_2)-\mu_1\mu_2 \\ &= \mu_1\mu_2-\mu_1\mu_2 \\ &= 0\end{aligned}$$

Theorema 2.2.3 :

Jika $X_1, X_2, \ldots, X_n$ variabel random dan $Y = X_1 + X_2 + \ldots + X_n$, maka

$$\mathrm{Var}(Y) = \sum_{i=1}^{n}\mathrm{Var}(X_i) + \sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n}\mathrm{Cov}(X_i,X_j)$$

kemudian jika $X_1, X_2, \ldots, X_n$ saling independent maka :

$$\mathrm{Var}(Y) = \sum_{i=1}^{n}\mathrm{Var}(X_i)$$

Bukti :

$Y = X_1 + X_2 + \ldots + X_n$

$$\mathrm{Var}(Y) = E[Y - E(Y)]^2$$

$$= E\left[\left((X_1 + X_2 + \ldots + X_n) - E(X_1) - E(X_2) - \ldots E(X_n)\right)^2\right]$$

$$= E\left[\left(X_1 - E(X_1) + X_2 - E(X_2) + \ldots + X_n - E(X_n)\right)^2\right]$$

$$= E\left[\sum_{i=1}^{n}\left(X_i - E(X_i)\right)^2 + \sum_{\substack{i=1 \\ i \neq j}}^{n}\sum_{j=1}^{n}\left(X_i - E(X_i)\right)\left(X_j - E(X_j)\right)\right]$$

$$= \sum_{i=1}^{n} E\left[\left(X_i - E(X_i)\right)^2\right] + \sum_{\substack{i=1 \\ i \neq j}}^{n}\sum_{j=1}^{n} E\left[\left(X_i - E(X_i)\right)\left(X_j - E(X_j)\right)\right]$$

$$= \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{\substack{i=1 \\ i \neq j}}^{n}\sum_{j=1}^{n} \mathrm{Cov}(X_i, X_j)$$

Jika X_1, X_2, ..., X_n saling bebas, maka dari theorema 2.2.2 $\mathrm{Cov}(X_i,X_j) = 0$ sehingga :

$$\mathrm{Var}(Y) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

Dalam sebuah kotak ada N bola yang diberi nomor 1 sampai N. Satu per satu bola diambil sampai ke-n pengambilan dengan $n \leq N$. Diasumsikan bahwa pengambilan dilakukan secara random, tanpa pengembalian dan untuk setiap bola mempunyai kesempatan yang sama. Jika Y merupakan jumlah bilangan (nomor) dari n pengambilan bola, maka untuk mengetahui mean dan varian dari Y dapat dianggap bahwa Y merupakan jumlah dari variabel random $X_1, X_2, \ldots, X_i, \ldots, X_n$.

Dimana X_i merupakan nomor pengambilan bola ke-i, untuk i = 1, 2, ..., n yang dipilih secara acak dan berdistribusi sama. Untuk menentukan mean dan varian Y tersebut digunakan theorema sebagai berikut :

Theorema 2.2. 4 :

Jika X jumlah dari n variabel random yang dipilih secara acak tanpa pengembalian dari N data (1 sampai N) maka mean dan varian dari X diberikan dengan :

$$E(X) = \frac{n(N+1)}{2} \text{ dan } Var(X) = \frac{n(N+1)(N-n)}{12}$$

Bukti :

Berdasarkan theorema 2.2.1 diperoleh bahwa jika $Y = X_1+X_2+\ldots+X_n$ maka $E(Y) = E(X_1)+E(X_2)+\ldots+E(X_n)$, sehingga jika X=Y maka $E(X) = E(X_1) + E(X_2) + \ldots + X_i + \ldots + X_n$. Karena $X_1, X_2, \ldots, X_i, \ldots, X_n$ dipilih secara random dan berdistribusi sama, maka $E(X) = nE(X_i)$ dimana X_i merupakan nomor pengambilan data ke-i. Peluang terambilnya data ke-i (X_i) adalah sama sebesar 1/N, yang dapat ditulis dengan $P(X_i = k) = 1/N$ untuk k=1, 2, ...,N, sehingga :

$$E(X_i) = \sum_{k=1}^{N} kf(k)$$

$$= \sum_{k=1}^{N} k.1/N$$

$$= \frac{1}{N}\frac{N(N+1)}{2}$$

$$= \frac{N+1}{2}$$

Jadi $E(X) = n.\ E(X_i) = \dfrac{n(N+1)}{2}$

Var(X)= ....?

Dari theorema 2.2.3 jika $Y = X_1 + X_2 + \ldots + X_n$ maka Var(Y)= $\sum_{i=1}^{n} \mathrm{var}(X_i) +$

$\sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} \mathrm{Cov}(X_i . X_j)$.

$$\mathrm{Var}(X_i) = E\left(X_i^2\right) - \left(E(X_i)\right)^2$$

$$\begin{aligned} &= \sum_{k=1}^{N} k^2 f(k) - \left(\frac{N+1}{2}\right)^2 \\ &= \sum_{k=1}^{N} k^2 1/N - \frac{(N+1)^2}{4} \qquad\qquad (*) \\ &= 1/N \frac{N(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4} \\ &= \frac{(N+1)(N-1)}{12} \end{aligned}$$

$\mathrm{Cov}(X_i, X_j) = E\left[(X_i - \mu_i)(X_j - \mu_j)\right]$ dimana mean X_j dan X_i =(N+1)/2, maka

$$\mathrm{Cov}(X_i, X_j) = E\left[(X_i - \mu_i)(X_j - \mu_j)\right]$$

$$\begin{aligned} &= \sum_{k=1}^{N} \sum_{\substack{s=1 \\ k \neq s}}^{N} \left(k - \frac{N+1}{2}\right)\left(s - \frac{N+1}{2}\right) f(x_i, x_j) \\ &= \sum_{k=1}^{N} \sum_{\substack{s=1 \\ k \neq s}}^{N} \left(k - \frac{N+1}{2}\right)\left(s - \frac{N+1}{2}\right) \frac{1}{N(N-1)} \end{aligned}$$

Akan diperluas untuk seluruh k dan s dari 1 sampai N, kecuali $k \neq s$ karena $X_i \neq X_j$ pada waktu yang sama. Yaitu dengan menjumlahkan dan menguranginya untuk k = s, sehingga covarian menjadi :

$$\mathrm{Cov}(X_i,X_j)=\sum_{k=1}^{N}\sum_{s=1}^{N}\left(k-\frac{N+1}{2}\right)\left(s-\frac{N+1}{2}\right)\frac{1}{N(N-1)}-\sum_{k=1}^{N}\left(k-\frac{N+1}{2}\right)^2\frac{1}{N(N-1)}$$

$$=\frac{1}{N(N-1)}\sum_{k=1}^{N}\left(k-\frac{N+1}{2}\right)\sum_{s=1}^{N}\left(s-\frac{N+1}{2}\right)-\frac{1}{N-1}\sum_{k=1}^{N}\left(k-\frac{N+1}{2}\right)^2\frac{1}{N}$$

(**)

Dengan catatan bahwa :

$$\sum_{i=1}^{N}\left(i-\frac{N+1}{2}\right)=\sum_{i=1}^{N}i-\sum_{i=1}^{N}\frac{N+1}{2}=\frac{N(N+1)}{2}-\frac{N(N+1)}{2}=0$$

Dari pers (*) dan (**) diperoleh :

$$\mathrm{Var}(X_i)=\sum_{k=1}^{N}\left(k-\frac{N+1}{2}\right)^2 1/N=\frac{(N+1)(N-1)}{12}$$

Sehingga $\mathrm{Cov}(X_i,X_j)=-\dfrac{(N+1)(N-1)}{12(N-1)}=-\dfrac{N+1}{12}$

Jadi :

$$\mathrm{Var}(X)=\sum_{i=1}^{n}\mathrm{Var}(X_i)+\sum_{i=1}^{n}\sum_{\substack{j=1\\ i\neq j}}^{n}\mathrm{Cov}(X_i,X_j)$$

$$=\frac{n(N+1)(N-1)}{12}+n(n-1)\left(-\frac{N+1}{12}\right)$$

$$=\frac{n(N+1)(N-n)}{12}$$

Definisi 2.2.10

Jika X variabel random, moment ke-r dari X dinotasikan dengan μ_r' yang didefinisikan $\mu_r'=E[X^r]$, dimana ekspektasi ada.

Definisi 2.2.11 :

Jika X variabel random, moment pusat ke-r dari X di a didefinisikan dengan $E[(X-a)^r]$, jika $a = \mu$ moment pusat dari X di μ dinotasikan dengan $\mu_r = E[(X-\mu)^r]$

Dengan cacatan bahwa $\mu_1 = E[(X-\mu)] = 0$ dan $\mu_2 = E[(X-\mu)^2]$ adalah varian X.

Definisi 2.2.12 :

Jika X variabel random dengan densitas f(.), nilai ekspektasi dari e^{tX} didefinisikan dengan fungsi pembangkit moment dari X. Jika nilai ekspektasi ada untuk setiap nilai t dalam interval $-h < t < h$, $h > 0$ maka fungsi pembangkit moment dinotasikan dengan $M_X(t)$ atau $M(t)$ adalah :

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x)dx \quad \text{jika X variabel random kontinu}$$

$$M_X(t) = E[e^{tX}] = \sum_{x} e^{tx} f(x) \quad \text{jika X variabel random diskrit}$$

Theorema 2.2.5 :

Jika fungsi pembangkit moment $M_X(t)$ dari variabel random X ada untuk $|t| \leq T$ (untuk suatu $T > 0$), maka $E[X^n]$ ada ($n = 1, 2, \ldots$) dan

$$E[X^n] = M_X^n(0) \equiv \frac{d^n}{dt^n} M_X(t)\Big|_{t=0}$$

Bukti :

$$M_X(t) = E\left[e^{tX}\right]$$

$$= E\left[1 + tX + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \ldots + \frac{(tX)^n}{n!} + \ldots\right]$$

$$= 1 + E[X]t + E\left[X^2\right]\frac{t^2}{2!} + E\left[X^3\right]\frac{t^3}{3!} + \ldots + E\left[X^n\right]\frac{t^n}{n!} + \ldots$$

Jika kedua ruas diturunkan terhadap t dan kemudian dimasukan t = 0, maka diperoleh :

$$M'_X(0) = \left(E[X] + E\left[X^2\right]t + E\left[X^3\right]\frac{t^2}{2!} + \ldots + E\left[X^n\right]\frac{t^{n-1}}{(n-1)!} + \ldots\right)\Bigg|_{t=0}$$

$$M'_X(0) = E[X]$$

sehingga pernyataan dalam theorema terbukti untuk kasus n =1. Hasil untuk n = 2, 3, ... diperoleh dengan cara yang sama yaitu dengan mencari turunan selanjutnya dan kemudian memasukan nilai t = 0, sehingga diperoleh :

$$M''_X(0) = E\left[X^2\right]$$

$$\vdots \qquad \vdots$$

$$M^n_X(0) = E\left[X^n\right]$$

jadi terbukti bahwa : $E[X^n] = M^n_X(0) \equiv \frac{d^n}{dt^n} M_X(t)\Big|_{t=0}$

2.3 PENGUJIAN HIPOTESIS

2.3.1 ERROR TIPE I DAN ERROR TIPE II

Hipotesis adalah asumsi atau dugaan mengenai sesuatu hal. Jika asumsi itu dikhususkan mengenai populasi, umumnya mengenai nilai-nilai parameter populasi atau mengenai sampel yang diambil dari populasi maka hipotesis itu disebut hipotesis statistik. Setiap hipotesis bisa benar atau salah, dan karenanya perlu diadakan penelitian sebelum hipotesis diterima atau ditolak. Langkah atau prosedur untuk menentukan apakah menerima atau menolak hipotesis disebut pengujian hipotesis. Dalam pengujian hipotesis ini digunakan dua hipotesis statistik yaitu hipotesis nol (H_o) dan hipotesis alternatif (H_1). Hipotesis nol merupakan hipotesis yang akan diuji oleh peneliti. Sehingga untuk memutuskan apakah H_o itu diterima atau ditolak dibutuhkan suatu kriteria tertentu untuk memutuskannya, diantaranya tingkat signifikansi (α). Dalam pengambilan kesimpulan H_o, kadang-kadang dilakukan dua tipe kesalahan yang mungkin dilakukan yaitu kesalahan tipe I (error tipe I) dan kesalahan tipe II (error tipe II).

Definisi 2.3.1 :

1. Menolak hipotesis nol yang seharusnya diterima disebut error tipe I, peluang dari error tipe I dinotasikan dengan α.
2. Menerima hipotesis nol ketika H_o salah disebut error tipe II, peluang dari error tipe II dinotasikan dengan β.

Definisi 2.3.2 :

Tingkat signifikansi (α) adalah peluang (probabilitas) error tipe I maximum dalam menolak H_o yang benar.

Tingkat signifikansi ini dapat diambil dengan asumsi bahwa H_o benar, dan terkadang disebut ukuran daerah kritis (daerah penolakan). Jika H_o benar peluang maximum untuk menolaknya adalah α dan peluang untuk menerimanya adalah $1-\alpha$.

Dalam pengambilan kesimpulan statistik ada kemungkinan untuk berbuat satu diantara dua tipe kesalahan. Maka dari itu peneliti harus dapat mencapai nilai kompromi yang merupakan keseimbangan yang optimal antara peluang-peluang yang diperbuat kedua tipe kesalahan itu. Untuk mencapai keseimbangan itu, maka digunakan fungsi kekuatan (power) yang didefinisikan sebagai peluang untuk menolak H_o ketika H_o salah, dengan peluang 1 - β.

2.4 DISTRIBUSI NORMAL, CHI-KUADRAT DAN THEOREMA LIMIT PUSAT

2.4.1 DISTRIBUSI NORMAL

Definisi 2.4.1 :

Variabel random X didefinisikan berdistribusi normal jika fungsi densitasnya adalah :

$f(x) = f(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$, dimana parameter μ dan σ adalah $-\sim<\mu<\sim$ dan $\sigma > 0$.

Definisi 2.4.2 :

Distribusi normal dengan $\mu = 0$ dan $\sigma = 1$ disebut distribusi normal standart.

Theorema 2.4.1 :

Jika X variabel random normal, maka $E(X) = \mu$, $Var(X) = \sigma^2$ dan

$M_X(t) = e^{\mu t + \sigma^2 t^2/2}$

Bukti :

Untuk membuktikannya akan digunakan fungsi pembangkit moment (fpm)

$$Mx(t) = E\left[e^{tX}\right]$$

$$= \int_{-\infty}^{\infty} e^{tx} \frac{-1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left(x^2-2\mu x+\mu^2-2tx\sigma^2\right)} dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left(\left[x-(\mu+t\sigma^2)\right]^2-(\mu+t\sigma^2)^2+\mu^2\right)} dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left(\left[x-(\mu+t\sigma^2)\right]^2-(2t\mu\sigma^2+t^2\sigma^4)\right)} dx$$

$$= e^{t\mu+\frac{1}{2}t^2\sigma^2} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2\sigma^2}\left(x-(\mu+t\sigma^2)\right)^2} dx$$

substitusi $y = \dfrac{x-(\mu+t\sigma^2)}{\sigma}$ maka $dx = \sigma\, dy$, sehingga :

$$M_x(t) = e^{t\mu+\frac{1}{2}t^2\sigma^2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}y^2} dy$$

$$M_x(t) = e^{t\mu+\frac{1}{2}t^2\sigma^2}$$

jadi $M_X(t) = e^{t\mu+\frac{1}{2}t^2\sigma^2}$

Berdasarkan theorema 2.2.5 diperoleh bahwa :

$$M_X(t) = e^{t\mu + \frac{1}{2}t^2\sigma^2}$$

$$M'_X(t) = e^{t\mu + \frac{1}{2}t^2\sigma^2}(\mu + t\sigma^2)$$

$$M'_X(0) = \mu$$

$$M''_X(t) = e^{t\mu + \frac{1}{2}t^2\sigma^2}(\mu + t\sigma^2)(\mu + t\sigma^2) + e^{t\mu + \frac{1}{2}t^2\sigma^2}.\sigma^2$$

$$M''_X(0) = \mu^2 + \sigma^2$$

karena

$$E(X) = M'_X(0) = \mu$$

$$Var(X) = M''_X(0) - \left(M'_X(0)\right)^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2$$

Jadi terbukti bahwa :$E(X) = \mu$, $Var(X) = \sigma^2$ dan $M_X(t) = e^{t\mu + \frac{1}{2}t^2\sigma^2}$

2.4.2 DISTRIBUSI CHI-KUADRAT

Definisi 2.4.3 :

Jika X variabel random dengan densitas

$f(x) = \frac{1}{\Gamma(k/2)}(1/2)^{k/2} x^{\frac{k}{2}-1} e^{\frac{-1}{2}x}$, maka X disebut berdistribusi chi-kuadrat dengan derajat bebas k.

Fungsi densitas chi-kuadrat merupakan kasus khusus dari fungsi densitas gamma dengan parameter r dan λ sama dengan k/2 dan ½. Dimana fungsi densitas gamma dengan parameter r dan λ adalah : $f(x; r, \lambda) = \frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1} e^{-\lambda x}$, dimana $r > 0, \lambda > 0$.

Jika X variabel random berdistribusi chi-kuadrat dengan derajat bebas k, maka :

$E(X) = k$, $Var(X) = 2k$ dan $M_X(t) = (1 - 2t)^{-k/2}$, $t < ½$.

Theorema 2.4.2 :

Jika variabel random $X_1, X_2, \ldots, X_k$ bebas dan identik berdistribusi normal standart dan jika Y merupakan jumlah dari kuadrat X_i, maka Y berdistribusi chi-kuadrat dengan derajat bebas k.

Bukti :

X_i iid ~ N(0,1)

$$f(x_i) = \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}x_i^2}$$

$$M_{Xi}(t) = e^{\frac{1}{2}t^2}$$

$X_i^2 \sim ?$

$$M_{Xi}^2(t) = E\left[e^{tX_i^2}\right]$$

$$= \int_{-\infty}^{\infty} e^{tx_i^2} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}x_i^2} dx_i$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}\left((1-2t)^{1/2}x_i\right)^2} dx_i$$

substitusi $y_i = x_i(1-2t)$ maka $dx_i = (1-2t)^{-1/2}\, dy_i$, sehingga :

$$M_{x_i}^2(t) = (1-2t)^{-1/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-1}{2}y^2} dy_i$$

$$M_{x_i}^2(t) = (1-2t)^{-1/2}$$

merupakan fungsi pembangkit moment untuk distribusi chi-kuadrat dengan derajat bebas 1. Jadi $X_i^2 \sim \chi^2_{db=1}$.

Misalkan $Y = \sum_{i=1}^{k} X_i^2$, maka :

$$M_Y(t) = M_{\sum\limits_{i=1}^{k} x_i^2}(t) = E\left[e^{t\sum\limits_{i=1}^{k} x_i^2}\right] = \prod_{i=1}^{k} E\left[e^{tx_i^2}\right]$$

$$= \prod_{i=1}^{k}(1-2t)^{-1/2}$$

$$= (1-2t)^{-k/2}, t < 1/2$$

merupakan fungsi pembangkit moment untuk distribusi chi-kuadrat dengan derajat bebas k. Jadi terbukti bahwa $Y = \sum\limits_{i=1}^{k} X_i^2 \sim \chi^2_{db=k}$

2.4.3 THEOREMA LIMIT PUSAT

Theorema 2.4.3 :

Misalkan diketahui f(.) fungsi densitas dengan mean μ dan varian terbatas σ^2. $\bar{X}_n$ adalah mean sampel dari sampel random berukuran n dari f(.). Variabel random

$$Z_n = \frac{\bar{X}_n - E\left[\bar{X}_n\right]}{\sqrt{Var\left(\bar{X}_n\right)}} = \frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}}$$

maka distribusi Z_n mendekati normal standart, untuk $n \to \sim$.

Bukti:

Untuk membuktikan theorema tersebut akan digunakan fungsi pembangkit moment (fpm). Sebab, jika fpm ada maka akan terjadi korespondensi satu-satu antara fpm dengan fungsi distribusi. Sehingga jika ada dua variabel random misal

X dan Y dengan fpm $M_X(t)$ dan $M_Y(t)$ keduanya ada dan sama, maka keduanya mempunyai fungsi distribusi yang sama.

Diketahui fpm dari distribusi normal standart adalah $M_X(t) = e^{\frac{1}{2}t^2}$.

Maka akan ditunjukkan bahwa $M_{Z_n}(t) = e^{\frac{1}{2}t^2}$.

$$M_{Z_n}(t) = E\left[e^{tZ_n}\right] = E\left[\exp tZ_n\right]$$

$$= E\left[\exp\left(t\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)\right] = E\left[\exp\left(\frac{t}{n}\sum\frac{X_i-\mu}{\sigma/\sqrt{n}}\right)\right]$$

$$= E\left[\prod_{i=1}^{n}\exp\frac{t(X_i-\mu)}{n\,\sigma/\sqrt{n}}\right] = \prod_{i=1}^{n}E\left[\exp\left(t\frac{X_i-\mu}{\sigma\sqrt{n}}\right)\right]$$

Jika $Y_i = X_i - \mu$ dan $M_Y\left(t/\sigma\sqrt{n}\right)$ menotasikan $M_{Y_i}\left(t/\sigma\sqrt{n}\right)$ maka :

$$M_{Z_n}(t) = \prod_{i=1}^{n}E\left[\exp\left(\frac{tY_i}{\sigma\sqrt{n}}\right)\right] = \prod_{i=1}^{n}M_{Y_i}\left(t/\sigma\sqrt{n}\right) = \left[M_Y\left(t/\sigma\sqrt{n}\right)\right]^n$$

Dimana :

$$M_Y\left(t/\sigma\sqrt{n}\right) = E\left[e^{Y\,t/\sigma\sqrt{n}}\right]$$

$$M_Y\left(t/\sigma\sqrt{n}\right) = E\left[1+\frac{Yt}{\sigma\sqrt{n}}+\frac{Y^2}{2!}\left(t/\sigma\sqrt{n}\right)^2+\frac{Y^3}{3!}\left(t/\sigma\sqrt{n}\right)^3+\ldots\right]$$

$$= 1+E[Y]\frac{t}{\sigma\sqrt{n}}+E\left[Y^2\right]\frac{t^2}{2!\sigma^2 n}+E\left[Y^3\right]\frac{t^3}{3!\sigma^3 n\sqrt{n}}+\ldots$$

$$= 1+E[X-\mu]\frac{t}{\sigma\sqrt{n}}+E\left[(X-\mu)^2\right]\frac{t^2}{2!\sigma^2 n}+E\left[(X-\mu)^3\right]\frac{t^3}{3!\sigma^3 n\sqrt{n}}+\ldots$$

Berdasarkan definisi 2.2.11 yang menyatakan bahwa $\mu_r = E\left[(X-\mu)^r\right]$, maka :

$$M_Y\left(t/\sigma\sqrt{n}\right) = 1 + \mu_1 \frac{t}{\sigma\sqrt{n}} + \mu_2 \frac{t^2}{2!\sigma^2 n} + \mu_3 \frac{t^3}{3!\sigma^3 n\sqrt{n}} + \ldots$$

Karena $\mu_1 = 0$ dan $\mu_2 = \sigma^2$, maka dapat dituliskan sebagai berikut :

$$M_Y\left(t/\sigma\sqrt{n}\right) = 1 + \frac{t^2}{2n} + \frac{\mu_3 t^3}{3!\sigma^3 n\sqrt{n}} + \ldots$$

$$= 1 + \frac{1}{n}\left(\frac{t^2}{2} + \frac{\mu_3 t^3}{3!\sigma^3\sqrt{n}} + \ldots\right)$$

Jika $u = \frac{t^2}{2} + \frac{\mu_3 t^3}{3!\sigma^3\sqrt{n}} + \ldots$ Maka $\left[M_Y\left(t/\sigma\sqrt{n}\right)\right]^n = \left[1 + \frac{u}{n}\right]^n$

Karena $\lim_{n\to\sim}\left[1 + \frac{u}{n}\right]^n = e^u = e^{\frac{1}{2}t^2}$, maka $\lim_{n\to\sim} M_{Z_n}(t) = e^{\frac{1}{2}t^2}$.

Maka dapat disimpulkan bahwa fungsi pembangkit moment dari Z_n adalah

$M_{Z_n}(t) = e^{\frac{1}{2}t^2}$, sehingga Z_n mendekati distribusi normal standart.