BAB II

TEORI PENUNJANG

2.3. VARIABEL RANDOM DAN FUNGSI DISTRIBUSI KUMULATIF

DEFINISI 1 (VARIABEL RANDOM)

Diberikan sebuah ruang probabilitas (Ω, A,P[.]), sebuah variabel random dinyatakan dengan X atau X(.), adalah sebuah fungsi dengan domain Ω dan co domain bilangan riil. Fungsi X(.) harus sedemikian sehingga kumpulan Ar, dimana kumpulan $Ar = \{\omega : X(\omega) \leq r\}$, melalui A untuk setiap bilangan riil r (Mood,53).

Contoh: Pelemparan sebuah koin

X mendefinisikan gambar

$\Omega = \{gambar, angka\}$

$X(\omega) = 1$ jika ω gambar

$= 0$ jika angka

maka variabel random X merupakan fungsi yang nilainya merupakan kumpulan bilangan riil yang dihasilkan dari eksperimen.

A = {$\varnothing$, gambar, angka dan Ω}.

Jika $r<0$, $\{\omega : X(\omega) \leq r\} = \varnothing$

$0 \leq r < 1$, $\{\omega : X(\omega) \leq r\}$ = angka

$r \geq 1$, $\{\omega : X(\omega) \leq r\}$ = gambar, angka = Ω

sehingga X(.) variabel random, karena untuk setiap r, kumpulan $\{\omega: X(\omega) \leq r\}$melalui A.

DEFINISI 2(FUNGSI DISTRIBUSI KUMULATIF)

Fungsi distribusi kumulatif dari sebuah variabel random X dinyatakan dengan $F_X(.)$ didefinisikan sebagai sebuah fungsi dengan domain bilangan riil dan co domain interval [0,1] yang memenuhi $F_X(x)=P[X \leq x]= P[\{\omega: X(\omega) \leq x\}]$ untuk setiap bilangan riil x (Mood, hal 54)

2.2 FUNGSI DENSITAS

DEFINISI 3 (VARIABEL RANDOM DISKRIT)

Sebuah variabel random X dikatakan diskrit, jika range X terhitung.

Jika sebuah variabel random X diskrit, maka cumulatif distribusi fungsi(cdf) yang berhubungan $\{F_X(.)\}$ akan terdefinisi diskrit. (Mood, hal 57).

DEFINISI 4 (FUNGSI DENSITAS DISKRET DARI SEBUAH VARIABEL RANDOM DISKRIT)

Jika X adalah sebuah variabel random diskrit dengan nilai $x_1, x_2, x_3, \ldots, x_n$ yang berbeda, maka fungsi yang didefinisikan dengan $f_X(.)$ merupakan fungsi densitas diskrit dari X dan didefinisikan sebagai:

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

DEFINISI 5 (VARIABEL RANDOM KONTINUE)

X disebut variabel random kontinue jika ada sebuah fungsi $f_X(x)$ sedemikian hingga

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\,du \quad \text{untuk setiap bilangan riil x.}$$

Cdf $F_X(.)$ dari sebuah variabel random kontinue X disebut kontinue mutlak.

DEFINISI 6 (FUNGSI DENSITAS PROBABILITAS DARI SEBUAH VARIABEL RANDOM KONTINUE)

Jika X adalah sebuah variabel random kontinue, maka fungsi $f_X(.)$ dalam

$$F_X(u) = \int_{-\infty}^{x} f_X(u)\,du$$

disebut fungsi densitas probabilitas dari X.

DEFINISI 7 (FUNGSI DENSITAS PROBABILITAS)

Fungsi-fungsi f(.) dengan domain bilangan riil dan codomain $[0,\infty)$ disebut fungsi densitas probabilitas, jika dan hanya jika:

(i) $f(x) \geq 0 \;\; \forall x.$

(ii) $\int_{-\infty}^{\infty} f(x)\,dx = 1$

DEFINISI 8 (FUNGSI PEMBANGKIT MOMENT)

Diambil X sebuah variabel random dengan densitas $f_X(.)$. Nilai harapan exp (tx) didefinisikan sebagai fungsi pembangkit moment dari X jika nilai harapan ada untuk semua nilai dari t dalam interval $-h < t < h$; $h>0$.

Fungsi pembangkit moment dinotasikan dengan $m_X(t)$ atau m(t) yaitu:

$$m(t)=E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} f_x(x)dx, \text{ jika variabel random X kontinue}$$

$$= \sum e^{tx} f_x(x) \quad \text{, jika variabel random X diskret.}$$

DEFINISI 9 (DISTRIBUSI UNIFORM)

Jika fungsi densitas probabilitas dari sebuah variabel random X diberikan dengan

$$f_X(x) = f_X(x;a,b) = \frac{1}{b-a}$$

Dimana parameter a dan b memenuhi $-\infty < a < b < x < \infty$, maka variabel random X didefinisikan terdistribusi uniform melalui interval [a,b], dan disebut distribusi uniform.

TEOREMA 2.1

Jika X terdistribusi uniform melalui [a,b], maka:

$$E[X] = \frac{a+b}{2}$$

$$Var[X] = \frac{(b-a)^2}{12}$$

$$M_x(t) = \frac{e^{bt} - e^{at}}{(b-a)t}$$

Bukti:

$$E(X) = \int_a^b x \frac{1}{b-a} dx$$

$$= \frac{b^2 - a^2}{2(b-a)}$$

$$= \frac{a+b}{2}$$

$$\text{Var}[X] = E[X^2] - [E[x]]^2$$

$$= \int_a^b x^2 \frac{1}{b-a} dx - \left(\frac{a+b}{2}\right)^2$$

$$= \frac{b^3 - a^3}{3(b-a)} - \frac{(a+b)^2}{4}$$

$$= \frac{(b-a)^2}{12}$$

$$m_X(t) = \frac{e^{bt} - e^{at}}{(b-a)t}$$

$$= E[e^{tx}]$$

$$= \int_a^b e^{tx} \frac{1}{b-a} dx$$

$$= \frac{e^{bt} - e^{at}}{(b-a)t}$$

DEFINISI 10 (DISTRIBUSI NORMAL)

Sebuah variabel random X didefinisikan terdistribusi normal jika diberikan densitas yaitu:

$$f_x(x) = \frac{1}{\tau\sqrt{2\pi}} \exp\left[-\frac{1}{2\tau^2}(x-\mu)^2\right]$$

dengan μ dan $\tau^2, -\infty < \mu < \infty, \tau > 0$

DEFINISI 11 (DISTRIBUSI EKSPONENSIAL)

Jika sebuah variabel random X mempunyai densitas:

$f_x(x, \lambda) = \lambda e^{-\lambda x}$ dengan $\lambda > 0$

maka X mempunyai distribusi eksponensial.

2.4. ESTIMASI DISTRIBUSI PADA ORDE STATISTIK

DEFINISI 12 (ORDE STATISTIK)

Ambil $X_1, X_2, X_3, \ldots, X_n$ menotasikan sebuah sampel random berukuran n dari sebuah kumulatif distribusi fungsi F(x). Maka $Y_1 \leq Y_2 \leq Y_3 \leq \ldots \leq Y_n$ dengan Y_i adalah X_i yang disusun agar increasing magnitude dan didefinisikan menjadi orde statistik berhubungan dengan sampel random $X_1, X_2, X_3, \ldots, X_n$ (Mood, hal 251).

Distribusi Empirik merupakan sebuah estimator alam dari sebuah fungsi distribusi F tidak diketahui dalam tidak adanya tambahan informasi mengenai bentuk dari F.

MLE menghasilkan suatu estimasi distribusi yang mempunyai nilai kerusakan monoton yang merupakan sebuah langkah untuk menganalisa data umur hidup dengan statistik.

MLE memberikan sebuah distribusi empirik pada orde statistik yaitu

$$F_n(x_{i:n}) = \sum_{j=1}^{i} f_n(x_{j:n}) w_j = \frac{i}{n} \qquad 1 \le i \le n$$

dengan $w_j = x_{j:n} - x_{j-1:n}$

$$f_n(x_{i:n}) = \frac{1}{n(x_{i:n} - x_{i-1:n})}$$

yang merupakan fungsi jumlah kumulatif $F_n(x_{i:n})$.

DEFINISI 13 (DENSITAS FUNGSI KERUSAKAN MONOTON)

Sebuah distribusi F yang berhubungan dengan distribusi umur hidup yang mempunyai densitas f mempunyai fungsi kerusakan monoton r(t) :

$$r(t) = \frac{f(t)}{1 - F(t)}$$

menggambarkan proses dari pemakaian.

Sebuah fungsi f terdefinisi pada interval $I \subseteq R$ ke R dikatakan increasing pada I jika $x \le x'$ dengan $x, x' \in I$ maka $f(x) \le f(x')$.

Sebaliknya fungsi f terdefinisi pada interval $I \subseteq R$ ke R dikatakan decreasing pada I jika $x \ge x'$ dengan $x, x' \in I$ maka $f(x) \ge f(x')$. Jika fungsi increasing atau decreasing pada sebuah interval maka dikatakan monoton pada interval itu.

DEFINISI 14 (FUNGSI IFR)

F IFR (Increasing Failure Rate) jika -log (1-F(x)) conveks pada F pada I nterval $(0,\infty)$. Jika F IFR, maka F kontinue mutlak kecuali pada akhir sisi kanan interval.

MLE sebagai suatu metode untuk mendapatkan estimator mendefinisikan transformasi fungsi jumlah kumulatif sebagai berikut :

$$H_F^{-1}(t) = \int_0^{F^{-1}(t)} [1 - F(u)]du, 0 \le t \le 1$$

$$H_F^{-1}(1) = \int_0^{F^{-1}(1)} [1 - F(u)]du = \mu$$

Untuk F IFR , F increasing kuat dalam intervalnya.

Interval F^{-1} tunggal terdefinisi pada (0,1).

$F^{-1}(1)$ diambil sama pada titik sisi kanan, mungkin $+\infty$.

$H_F^{-1}(0)=0$ dan

$$H_F^{-1}(1) = \int_0^{F^{-1}(t)} [1-F(u)]du = \mu$$

dengan μ adalah mean dari F. (Jika F IFR maka $\mu < \infty$). Oleh karena itu H_F adalah distribusi dengan batasan pada $[0, \mu]$ karena $H^{-1}{}_F$ (Invers H_F) adalah increasing kuat pada [0,1].

Jika $G(x)=1- e^{-x}$ untuk $x \ge 0$, maka $H_G^{-1}(t) = t$ dan H_G terdistribusi uniform pada [0,1].

$$G(x)=1-\exp(-x)$$

$$\text{Exp}(-x)=1-G(x)$$

$$-x = \ln[1-G(x)]$$

$$x = -\ln[1-G(x)]$$

$$G^{-1}(x) = -\ln[1-x]$$

$$H_G^{-1}(x) = \int_0^{-\ln[1-x]} \left[1 - \left(1 - e^{-u}\right)\right] du$$

$$= \int_0^{-\ln(1-x)} \exp(-u)\, du$$

$$= x$$

Kita tuliskan H^{-1} untuk $H^{-1}{}_F$.

Catatan bahwa

$$\frac{d}{dt} H_F^{-1}(t)\Big|_{t=F(x)} = \frac{1-F(x)}{f(t)} = \frac{1}{r(t)}$$

Jika F IFR, maka r increasing dan $H^{-1}{}_F$ (t) adalah sebuah fungsi konkav dari

$$\mu = \int_0^\infty x\, dF(x)$$

$$H_n^{-1}\left(\frac{i}{n}\right) \overset{def}{=} H_{Fn}^{-1}\left(\frac{i}{n}\right) = \int_0^{xi:n} [1 - F_n(u)]\, du$$

$t \in [0,1]$ atau H_F adalah sebuah distribusi convexs pada $[F^{-1}(0), \mu_0]$ dengan

$\mu = \int_0^\infty x\, dF(x)$.

Jika F digantikan F_n, maka kita dapatkan :

Untuk $1 \le i \le n$. $H^{-1}{}_n(t)$ untuk $0 \le t \le 1$ didapatkan dengan interpolasi linier

yaitu :

$H_n^{-1}(t) = nX_{1:n}t$ untuk $0 \le t \le 1/n$

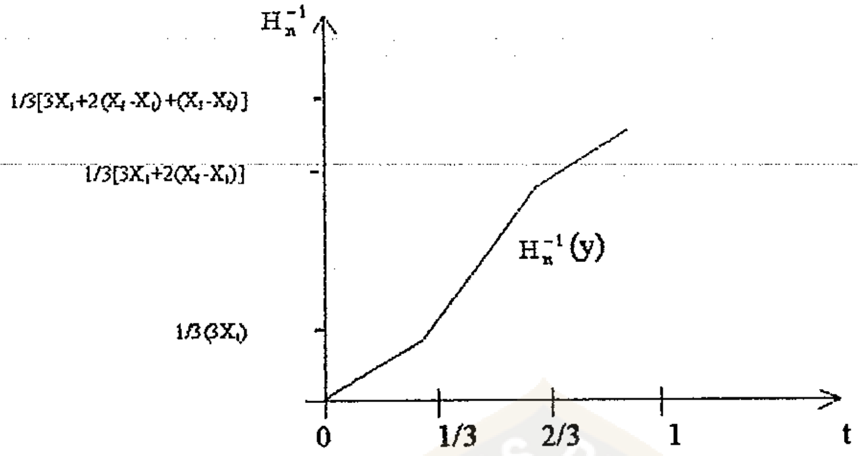
Sementara

$H_n^{-1}(t) = 1/n\, [\, nX_{1:n} + \ldots + (n-i+1)(X_{i:n} - X_{i-1:n}) + [t-i/n](n-i)\,(X_{i:n} - X_{i-1:n})$

Untuk $i/n \le t \le (i+1)/n$ dan $1 \le i \le n-1$

Grafik berikut menjelaskan bahwa H_n^{-1} (t) untuk n=3. H_n^{-1} kontinue dan

H_n^{-1} (1) adalah mean sampel .



Gambar 1.

Secara umum

H_n^{-1} (i/n) = 1/n [$nX_{1:n}$+(n-1)($X_{2:n}$-$X_{1:n}$) +…+ (n-i+1)($X_{i:n}$ – $X_{i-1:n}$)] untuk $1 \leq i \leq n-1$.

Persamaan dalam kurung siku menunjukkan total waktu pada uji pada observasi ke i. Jika n item ditempatkan pada uji hidup pada waktu ke nol, maka n item bertahan pada waktu $X_{1:n}$, n-1 item bertahan melalui interval ($X_{1:n}$, $X_{2:n}$), dst, sementara (n-i+1) bertahan pada interval [$X_{i-1:n}$, $X_{i:n}$]. Total waktu pada uji statistik adalah dasar dari penyelesaian permasalahan uji hidup (life testing).

2.5. UJI HIPOTESA

Ada dua wilayah besar dalam statistik inferensi yaitu estimasi parameter dan uji hipotesa.

DEFINISI 15 (HIPOTESA STATISTIK)

Sebuah hipotesa statistik adalah sebuah dugaan tentang distribusi dari satu atau lebih variabel random. Jika hipotesa statistik lengkap dengan distribusi spesifik(tertentu), maka disebut simpel, dan sebaliknya disebut hipotesa komposit.(Mood, hal 402).

DEFINISI 16 (UJI DARI SEBUAH HIPOTESA STATISTIK)

Sebuah uji dari sebuah hipotesa statistik H adalah sebuah prosedur untuk memutuskan apakah menolak H.(Mood, hal 403).

DEFINISI 17 (RANDOM TEST)

Sebuah uji(test) Y dari sebuah hipotesa H didefinisikan sebagai random test jika Y didefinisikan dengan:

$\Psi_Y(x_1, x_2, x_3, \ldots, x_n) = P[$H ditolak , dimana $(x_1, x_2, x_3, \ldots, x_n)$ diobservasi$]$.

Fungsi $\Psi_Y(x_1, x_2, x_3, \ldots, x_n)$ disebut fungsi kritis dari uji Y.(Mood, hal404)

DEFINISI 18 (KESALAHAN TYPE I DAN II)

Kesalahan type I: Menolak H_0 pada saat H_0 benar.

Kesalahan type II: Menerima H_0 pada saat H_0 salah (H_1 benar). (Mood, hal 404).

DEFINISI 19 (POWER FUNCTION)

Ambil Y sebuah uji hipotesa nol H_0. Power function dari uji Y dinotasikan dengan $\pi_Y(\theta)$, merupakan probabilitas bahwa H_0 ditolak disaat distribusi sampel mempunyai parameter θ. (Mood, hal 406).

TEORI ASYMPTOTIK

Teori asymptotik dari sebuah orde statistik berkenaan dengan distribusi dari $X_{i:n}$ yang merupakan standarisasi untuk $n \to \infty$. Hal pertama biasanya diasumsikan bahwa $X_{i:n}$ adalah orde statistik ke-i dalam sebuah sampel berukuran n dari beberapa populasi dengan cumulative distribution function(cdf) F(x).

TEOREMA 2.2 (TEOREMA LIMIT CENTRAL)

Ambil f(x) densitas dengan mean μ dan varian finite σ^2. Ambil $\overline{X}_n$ mean sampel dari sebuah sampel random berukuran n dari f(x). Ambil variabel random Z_n yang didefinisikan dengan:

$$Z_n = \frac{\overline{X}_n - E[\overline{X}_n]}{\sqrt{\mathrm{var}[\overline{X}_n]}} = \frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}}$$

Maka distribusi Z_n mendekati distribusi normal standard untuk $n \to \infty$.

Dari teorema dinyatakan bahwa distribusi pendekatan untuk Z_n adalah sebuah distribusi normal, atau menyatakan bahwa $\overline{X}_n$ itu sendiri asymptotik, terdistribusi seperti distribusi normal dengan mean μ dan varian $\tau^2/_n$.

Bukti

Teorema limit central dapat dibuktikan dengan suatu asumsi bahwa moment generating function(m g f) untuk mean sampel mendekati moment generating function(m g f) dari distribusi normal.

M g f dari distribusi normal standard adalah exp(1/2 t^2) = m(t)

$M_{Z_n}(t) = E[\exp(tZ_n)]$

Akan ditunjukkan bahwa $m_Z(t)$ mendekati m(t) pada n besar.

$$= E\left[\exp\left(t\frac{\overline{X}-\mu}{\tau/\sqrt{n}}\right)\right]$$

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

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

$$= \prod_{i=1}^{n} E\left[\exp\left(\frac{t}{\sqrt{n}}\frac{X_i-\mu}{\tau}\right)\right] \text{ karena independent.}$$

Ambil $Y_i = (X_i-\mu)/\tau$, maka $m_{Y_i}(t)$ m g f dari Y_i, independent karena semua Y_i mempunyai distribusi sama.

Ambil $m_Y(t) = m_{Y_i}(t)$ maka:

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

$$= \prod_{i=1}^{n} m_Y\left(\frac{t}{\sqrt{n}}\right)$$

$$= \left[m_Y\left(\frac{t}{\sqrt{n}}\right)\right]^n$$

oleh karena itu $m_z(t) = \left[m_Y\left(\frac{t}{\sqrt{n}}\right)\right]^n$

Derivatif ke-r dari $m_Y(t/\sqrt{n})$ dengan t=0 memberikan moment ke-r dari mean dengan densitas f(x) dibagi dengan $(\tau/\sqrt{n})^r$ sehingga dapat ditulis:

$$M_Y(t/\sqrt{n}) = 1 + \frac{\mu_1 t}{\tau\sqrt{n}} + ½!\,\frac{\mu_2}{\tau^2}\left(\frac{t}{\sqrt{n}}\right)^2 + 1/3!\,\frac{\mu_3}{\tau^3}\left(\frac{t}{\sqrt{n}}\right)^3 + \ldots..$$

Karena $\mu_1 = 0$, $\mu_2 = \tau^2$ dapat dituliskan

$$M_Y(t/\sqrt{n}) = 1 + 1/n\left(\frac{t^2}{2} + \frac{\mu_{3'}\tau^3}{3!\tau^3\sqrt{n}} + \ldots\right)$$

u = menunjukkan ekspresi dalam kurung.

maka $\lim_{n\to\infty} (1+u/n)^n = \exp(1/2t^2)$, sehingga $\lim_{n\to\infty} m_{zn}$ mempunyai moment generating function yang sama seperti normal standard .

Teorema berikut menunjukkan bahwa dengan moment generating function yang sama, maka dua variabel random akan memiliki fungsi distribusi yang sama.

TEOREMA 2.3 (KESAMAAN DISTRIBUSI FUNGSI DUA VARIABEL RANDOM)

Ambil X dan Y dua variabel random dengan densitas $f_X(x)$ dan $f_Y(y)$. $m_X(t)$ dan $m_Y(t)$ ada dan sama untuk semua t dalam interval $-h<t<h$ untuk beberapa $h>0$. Maka dua fungsi distribusi kumulatif $F_X(x)$ dan $F_Y(x)$ adalah sama.

Bukti

Teorema diatas akan dibuktikan dengan asumsi bahwa suatu fungsi distribusi akan didapatkan jika diketahui moment generating functionnya.

Misalnya:

X variabel random terdistribusi normal N(0,1)

$Y = X^2$. Tentukan distribusi dari Y.

$m_Y(t) = E[e^{tY}]$

$$= \int_{-\infty}^{\infty} e^{tx^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2(1-2t)}\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \frac{(1-2t)^{-\frac{1}{2}}}{(1-2t)^{-\frac{1}{2}}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2(1-2t)}\, dx$$

$$= (1-2t)^{-\frac{1}{2}}$$

$$= \left(\frac{1/2}{1/2 - t}\right)^{\frac{1}{2}} \quad \textit{untuk} \quad t < 1/2$$

merupakan moment generating function dari distribusi Gamma dengan parameter $r<1/2$, dan $\lambda=1/2$ (atau disebut juga distribusi chi-kuadrat dengan.derajat bebas 1.

Dengan demikian dapat disimpulkan bahwa pendekatan distribusi dari Z_n adalah distribusi normal standart atau X_n sendiri asymptotik terdistribusi seperti distribusi normal dengan mean μ dan varian τ^2/n.