PROBABILITAS (MAS106)

Ekspektasi, Variansi, dan MGF Variabel Random Diskrit

Tim Dosen Pengajar Probabilitas

OUTLINE

- Ekspektasi Variabel Random Diskrit
- Variansi Variabel Random Diskrit
- MGF Variabel Random Diskrit

Nilai yang diekspektasikan (*Expected Value*)

Nilai yang diharapkan disebut juga dengan harapan, atau mean.

DISKRIT

Misalkan X adalah variabel acak diskrit mengambil nilai x_i dengan probabilitas $f(x_i)$, $i = 1, 2, \ldots, n$, maka ekspektasi dari X dilambangkan dengan E(X).

$$E(X) = \sum_{i=1}^{n} x_i f(x_i)$$

Jika X adalah variabel acak diskrit yang mengasumsikan nilai $x_1, x_2, \ldots$,(*infinite*) dengan probabilitas $f(x_i) = P(X = x_1), i = 1, 2, \ldots,$ maka nilai yang diekspektasikan dari X didefinisikan sebagai

$$E(X) = \sum_{i} x_i f(x_i) = \sum_{i} x_i P(X = x_i) \quad (1)$$

dengan ketentuan

$$\sum_{i} |x_i| \, P(X = x_i) < \infty \quad (2)$$

Notasi alternatif yang sering digunakan: EX atau $\mu(X)$ atau μ_X

DEFINISI

Sifat-Sifat Ekspektasi

Dari definisi ekspektasi dan sifat-sifat penjumlahan atau integral yang sudah dikenal, maka:

$$E(cX) = cE(X),$$
$$E(cX + d) = cE(X) + d,$$

di mana c dan d adalah konstanta.

$X \geq c$, menyiratkan $E(X) \geq c$, dan, khususnya, $X \geq 0$ menyiratkan $E(X) \geq 0$

Theorem 1.8.1. *Let X be a random variable and let $Y = g(X)$ for some function g.*

(a) *Suppose X is continuous with pdf $f_X(x)$. If $\int_{-\infty}^{\infty} |g(x)| f_X(x)\, dx < \infty$, then the expectation of Y exists and it is given by*

$$E(Y) = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx. \tag{1.8.2}$$

(b) *Suppose X is discrete with pmf $p_X(x)$. Suppose the support of X is denoted by $\mathcal{S}_X$. If $\sum_{x \in \mathcal{S}_X} |g(x)| p_X(x) < \infty$, then the expectation of Y exists and it is given by*

$$E(Y) = \sum_{x \in \mathcal{S}_X} g(x) p_X(x). \tag{1.8.3}$$

$$Y = g(X),$$

$$EY = \sum_{i=1}^{n} g(x_i) f(x_i) \quad \text{or} \quad EY = \sum_{i=1}^{\infty} g(x_i) f(x_i) \quad \text{or} \quad EY = \int_{-\infty}^{\infty} g(x) f(x)\, dx,$$

$$EX^k = \sum_{i=1}^{n} x_i^k f(x_i) \quad \text{or} \quad EX^k = \sum_{i=1}^{\infty} x_i^k f(x_i) \quad \text{or} \quad EX^k = \int_{-\infty}^{\infty} x^k f(x)\, dx.$$

Variansi

Variansi dari variabel acak X dinotasikan dengan Var(X) dan didefinisikan dengan

$$Var(X) = E(X - E(X))^2 \quad (5)$$

Notasi lainnya yang sering digunakan adalah $\sigma^2(X)$ dan σ_X^2

Persamaan (5) dapat dinyatakan dengan formulasi sebagai berikut:

$$Var(X) = E(X^2) - (E(X))^2 \quad (6)$$

DEFINISI

Sifat-Sifat Variansi

Dari persamaan (5), itu juga segera mengikuti bahwa

$$Var(cX) = c^2 Var(X),$$
$$Var(cX + d) = c^2 Var(X),$$

di mana c dan d adalah konstanta.

$$Var(c) = Var(d) = 0$$

Contoh

X adalah jumlah lemparan di mana Gambar pertama muncul. Misalkan Y menjadi jumlah lemparan sebelum Gambar pertama. Maka Y = X-1. Dalam hal ini, fungsi g adalah g(x) = x−1, yang inversnya diberikan oleh $g^{-1}(y)$ = y+1. Ruang dari Y adalah D_Y = {0, 1, 2, . . .}. Fungsi probabilitas (pmf) dari X dan pmf dari Y adalah

$$P(X=x)=\left(\frac{1}{2}\right)^{x-1}\left(\frac{1}{2}\right)=\left(\frac{1}{2}\right)^{x}, \qquad x=1,2,3,\ldots$$

$$P_Y(y)=P_X(y+1)=\left(\frac{1}{2}\right)^{y+1}, \qquad untuk\ y=0,1,2,\ldots$$

Misalkan $Z=e^{-Y}$. Karena $(2e)^{-1}<1$, kita memiliki teorema 1.8.1 bahwa

$$E(Z)=E(e^{-Y})=\sum_{y=0}^{\infty} e^{-y}\left(\frac{1}{2}\right)^{y+1}$$

$$=e\sum_{y=0}^{\infty}\left(\frac{1}{2}e^{-1}\right)^{y+1}=\frac{e}{1-(1/(2e))}=\frac{2e^2}{2e-1}$$

Latihan

Misalkan X memiliki pmf

$$p(x) = \begin{cases} \frac{x}{6} & x = 1, 2, 3 \\ 0 & lainnya \end{cases}$$

Hitung $E(6X^3 + X)$ dan $\mathrm{Var}(-6X)$

Fungsi Pembangkit Momen

Fungsi $M(t) = E(e^{tX})$, didefinisikan untuk semua *t* dalam R untuk $E(e^{tX})$ berhingga (*finite*), disebut dengan *moment generating function* (mgf) dari X.

Biarkan X menjadi variabel acak sedemikian rupa sehingga untuk beberapa $h > 0$, ekspektasi dari e^{tX} ada untuk $-h < t < h$. Fungsi pembangkit momen dari X didefinisikan sebagai fungsi $M(t) = E(e^{tX})$, untuk $-h < t < h$. Kami menggunakan singkatan mgf untuk menunjukkan fungsi pembangkit momen dari variabel acak.

mgf ada di lingkungan terbuka 0, mencakup interval bentuk (−h, h) untuk beberapa $h > 0$. Selanjutnya, terbukti bahwa jika kita menetapkan t = 0, kita memiliki M(0) = 1. Agar mgf ada, harus ada dalam interval terbuka sekitar 0.

DEFINISI

Contoh

Misalkan kita memiliki spinner yang adil dengan angka 1, 2, dan 3 di atasnya. Biarkan X menjadi jumlah putaran sampai angka 3 yang pertama terjadi. Dengan asumsi bahwa putarannya bebas, pmf dari X adalah

$$p(x) = \frac{1}{3}\left(\frac{2}{3}\right)^{x-1}, \qquad x = 1, 2, 3, \ldots$$

Lalu, menggunakan deret geometri, mgf dari X adalah

$$M(t) = E(e^{tX}) = \sum_{x=1}^{\infty} e^{tx}\frac{1}{3}\left(\frac{2}{3}\right)^{x-1} = \frac{1}{3}e^{t}\sum_{x=1}^{\infty}\left(e^{t}\frac{2}{3}\right)^{x-1} = \frac{1}{3}e^{t}\left(1 - e^{t}\frac{2}{3}\right)^{-1}$$

asalkan $e^{t}\left(\frac{2}{3}\right) < 1$; yaitu, $t < \log(3/2)$. Interval terakhir ini adalah interval terbuka 0; karenanya, mgf dari X ada.

Theorem 1.9.2. *Let X and Y be random variables with moment generating functions M_X and M_Y, respectively, existing in open intervals about 0. Then $F_X(z) = F_Y(z)$ for all $z \in R$ if and only if $M_X(t) = M_Y(t)$ for all $t \in (-h, h)$ for some $h > 0$.*

Karena distribusi yang memiliki mgf $M(t)$ sepenuhnya ditentukan oleh $M(t)$, tidak mengherankan jika kita dapat memperoleh beberapa sifat distribusi langsung dari $M(t)$. Misalnya, keberadaan $M(t)$ untuk $-h < t < h$ menyiratkan bahwa turunan dari $M(t)$ dari semua ordo ada pada $t = 0$. Juga, teorema dalam analisis memungkinkan kita untuk menukar orde diferensiasi dan integrasi (atau penjumlahan dalam kasus diskrit).

Sifat MGF

$$M_{cX}(t) = M_X(ct)$$
$$M_{cX+d}(t) = e^{dt} M_X(ct)$$

di mana c dan d adalah konstanta.

Turunan Pertama MGF

$$M'(t) = \frac{dM(t)}{dt} = \sum_x xe^{tx}p(x)$$

Setelah menetapkan t = 0, kami memiliki dalam kedua kasus

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

proof

$$\frac{d}{dt}M_X(t)\bigg|_{t=0} = \left(\frac{d}{dt}Ee^{tX}\right)\bigg|_{t=0} = E\left(\frac{d}{dt}e^{tX}\bigg|_{t=0}\right)$$

$$= E(Xe^{tX}|_{t=0}) = EX.$$

Turunan Kedua MGF

$$M''(t) = \frac{d^2M(t)}{dt^2} = \sum_x x^2 e^{tx} p(x)$$

sehingga

$$M''(0) = E(X^2).$$

Dengan demikian, Var(X) sama dengan

$$Var(X) = \sigma^2 = E(X^2) - [E(X)]^2 = M''(0) - [M'(0)]^2$$

Secara Umum

$$M^{(m)}(0) = E(X^m)$$

$$E(X^m) = \sum_x x^m p(x)$$

$$\left.\frac{d^m}{dt^m} M_X(t)\right|_{t=0} = E(X^m)$$

Jumlah semacam ini, dalam mekanika, disebut **momen**. Karena $M(t)$ menghasilkan nilai $E(X^m)$, m = 1, 2, 3, . . . , itu disebut fungsi pembangkit momen (mgf). Faktanya, kita terkadang menyebut $E(X^m)$ momen ke-m dari distribusi, atau momen ke-m dari X.

Dalam beberapa perhitungan, rumus berikut terbukti sangat berguna. Pertama, dalam menjumlahkan suku tak hingga dari deret geometri, kita mendapatkan:

$$\sum_{x=k}^{\infty} t^x = \frac{t^k}{1-t}, \quad k = 0, 1, \ldots, \text{ and } |t| < 1.$$

Next:

$$\sum_{x=1}^{\infty} x t^x = t \sum_{x=1}^{\infty} x t^{x-1} = t \sum_{x=1}^{\infty} \frac{d}{dt} t^x = t \frac{d}{dt} \sum_{x=1}^{\infty} t^x$$

$$= t \frac{d}{dt}\left(\frac{t}{1-t}\right) = \frac{t}{(1-t)^2}.$$

$$\sum_{x=2}^{\infty} x(x-1) t^x = t^2 \sum_{x=2}^{\infty} x(x-1) t^{x-2} = t^2 \sum_{x=2}^{\infty} \frac{d^2}{dt^2} t^x$$

$$= t^2 \frac{d^2}{dt^2} \sum_{x=2}^{\infty} t^x = t^2 \frac{d^2}{dt^2}\left(\frac{t^2}{1-t}\right) = \frac{2t^2}{(1-t)^3}.$$

Distribusi Bernoulli

Percobaan Bernoulli adalah eksperimen acak, yang hasilnya dapat diklasifikasikan hanya dalam satu dari dua cara yang saling eksklusif dan lengkap, misalnya, berhasil atau gagal.

Urutan percobaan Bernoulli terjadi ketika percobaan Bernoulli dilakukan beberapa kali independen sehingga probabilitas keberhasilan, misalnya p, tetap sama dari percobaan ke percobaan. Artinya, dalam urutan seperti itu, kami membiarkan p menunjukkan probabilitas keberhasilan pada setiap percobaan.

Biarkan X menjadi variabel acak yang terkait dengan percobaan Bernoulli :

X(sukses) = 1 dan X(gagal) = 0.

Artinya, dua hasil, sukses dan gagal, masing-masing dilambangkan dengan satu dan nol.

Fungsi probabilitas (pmf) dari X

$$p(x) = p^x(1-p)^{1-x}, \qquad x = 0, 1$$

dan X memiliki distribusi Bernoulli.

Ekspektasi dan Variansi

$$E(X) = p$$
$$Var(X) = p(1-p)$$

Distribusi Binomial

- Konsep percobaan binomial, yang dimaksudkan untuk menjadi percobaan yang menghasilkan dua kemungkinan hasil, satu disebut sukses, dilambangkan dengan *S* dan terjadi dengan probabilitas *p*, dan yang lainnya disebut kegagalan, dilambangkan dengan *F* dan terjadi dengan kemungkinan *q* = 1-*p*.
- Eksperimen binomial dilakukan n kali independen (dengan p tetap sama), dan misalkan X adalah variabel acak menunjukkan jumlah keberhasilan, x = 0, 1, . . . , n, dengan probabilitas masing-masing:

$$P(X=x) = f(x) = \binom{n}{x} p^x q^{n-x}, \qquad x = 0, 1, \ldots, n, 0 < p < 1, \ q = 1 - p.$$

$$X \sim B(n, p) \Rightarrow EX = np, \quad Var(X) = npq, \quad \text{and} \quad M_X(t) = (pe^t + q)^n, \qquad t \in \Re.$$

Contoh

Jika mgf dari variabel acak X adalah

$$M(t) = \left(\frac{2}{3} + \frac{1}{3}e^t\right)^5$$

Maka X memiliki distribusi Binomial dengan $n = 5$ dan $p = 1/3$, yaitu pmf dari X adalah

$$p(x) = \begin{cases} \binom{5}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{5-x} & x = 0, 1, 2, \dots, 5 \\ 0 & lainnya \end{cases}$$

$E(X) = np = 5/3$ dan $Var(X) = np(1-p) = 10/9$

Latihan

1. Diberikan X variabel acak dengan fungsi probabilitas

$$p(x) = p^x(1-p)^{1-x}, \qquad x = 0, 1$$

Tentukan

i) $E(X)$, $E(X-p)$, dan $E(X-2p)$

ii) $Var(X)$, $Var(X+p)$, dan $Var(X-p)$

2. Misalkan $X \sim B(2, p)$, hitung $E(X)$, $E(X/2)$, $\mathrm{Var}(X)$, dan $\mathrm{Var}(X/2)$

NEXT

- Jabarkan mendapatkan mgf dari distribusi Bernoulli dan Binomial
- Dapatkan E(X) dan Var(X) dari mgf yang didapatkan untuk masing-masing distribusi

Distribusi Geometri

- Distribusi ini muncul dalam situasi percobaan binomial ketika percobaan dilakukan secara independen (dengan probabilitas *p* konstan dari S) sampai **S pertama terjadi**.
- Variabel acak X yang menunjukkan jumlah percobaan yang diperlukan adalah variabel acak yang terdistribusi secara geometris dengan parameter *p* dan distribusinya adalah distribusi geometrik dengan parameter *p*. X mengambil nilai 1, 2, . . . dengan probabilitas masing-masing:

$$P(X = x) = f(x) = pq^{x-1}, \qquad x = 1, 2, \ldots, \; 0 < p < 1, \; q = 1 - p. \qquad (*)$$

Jika X terdistribusi secara geometris dengan parameter *p*, maka

$$EX = \frac{1}{p}, \quad Var(X) = \frac{q}{p^2}, \quad M_X(t) = \frac{pe^t}{1 - qe^t}, \qquad t < -\log q.$$

Distribusi Binomial Negatif

- Percobaan Bernoulli independen, masing-masing dengan probabilitas p keberhasilan, dilakukan. Jika X adalah banyaknya percobaan yang diperlukan untuk mengamati keberhasilan ke-*s*, *s* = 2, 3,..., maka X disebut variabel acak binomial negatif dengan parameter *s* dan *p*; fungsi probabilitas untuk X adalah

$$P(X = x) = f(x) = \binom{x-1}{s-1} p^s (1-p)^{(x-s)}$$

dengan $EX = \frac{s}{p}$ $\quad Var(X) = \frac{s(1-p)}{p^2}$ $\quad M(t) = p^s(1-(1-p)e^t)^{-s}, \quad untuk\ t < -\log(1-p)$

- Distribusi binomial negatif adalah distribusi dari variabel acak X banyaknya percobaan sampai diperoleh **sukses yang ke-s (pada ulangan ke-x)**. Untuk distribusi binomial negatif, **banyaknya percobaan tetap** dan **banyaknya sukses random/ acak** (berkebalikan dengan distribusi binomial).

Distribusi Poisson

- Variabel acak X mengambil nilai 0, 1, . . . dengan masing-masing probabilitas yang diberikan pada (*) dikatakan memiliki distribusi Poisson dengan parameter λ; Distribusinya disebut distribusi Poisson dengan parameter λ.
- X terdistribusi Poisson dengan parameter λ dilambangkan dengan $X \sim P(\lambda)$.

$$P(X = x) = f(x) = e^{-\lambda}\frac{\lambda^x}{x!}, \qquad x = 0, 1, \ldots, \ \lambda > 0.$$

Jika X ~ P(λ) , maka

$$EX = \lambda, \quad Var(X) = \lambda, \quad \text{and} \quad M_X(t) = e^{\lambda e^t - \lambda}, \qquad t \in \Re.$$

- Kebakaran di kota tertentu terjadi menurut proses Poisson. Jika ada 10 kebakaran dalam satu minggu, berapa probabilitas bahwa setidaknya salah satu dari mereka terjadi pada hari Jumat?

Distribusi Hipergeometrik

- Distribusi ini cukup sering terjadi dan cocok untuk menggambarkan situasi jenis berikut: m objek identik (misalnya, bola) dicampur secara menyeluruh dengan n objek identik (yang lagi-lagi dapat dianggap sebagai bola) tetapi berbeda dari m objek. Dari m+n benda ini, r diambil **tanpa pengembalian**, dan misalkan X adalah bilangan di antara r yang berasal dari m benda.
- Variabel acak X mengambil nilai 0, 1, . . . , min(r, m) dengan probabilitas masing-masing diberikan di bawah ini. Sebenarnya, dengan mendefinisikan $\binom{m}{x} = 0$ untuk x > m, kita mendapatkan:

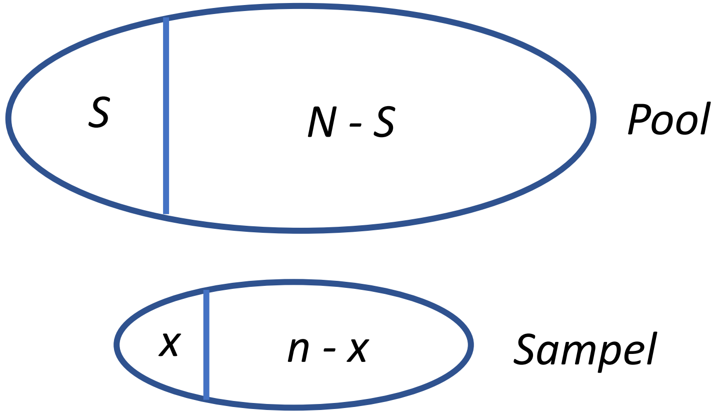
$$P(X = x) = f(x) = \frac{\binom{m}{x}\binom{n}{r-x}}{\binom{m+n}{r}}, \qquad x = 0, \ldots, r;$$

m dan *n* dapat disebut sebagai parameter distribusi.

Distribusi Hipergeometrik

Atau formula yang lain, sebagai berikut

$$P(X = x) = f(x) = \frac{\binom{S}{x}\binom{N-S}{n-x}}{\binom{N}{n}}, \qquad x = 0, 1, \dots, n$$



dengan pool ukuran *N* memuat *S* sukses dan (*N* – *S*) gagal, dan sampel acak berukuran *n* diambil dari pool, banyaknya sukses *X* dalam sampel mengikuti distribusi hipergeometri dengan parameter *n*, *S*, dan *N*.

Jika X~Hipergeometri(*N*, *S*, *n*) $\Rightarrow$ $EX = n\frac{S}{N}$ $\quad Var(X) = n\frac{S}{N}\frac{N-S}{N}\frac{N-n}{N-1}$

LATIHAN minggu lalu

Let X be a r.v. with p.d.f. $f(x) = e^{-x}$, $x > 0$. Then:

(i) Find the m.g.f. $M_X(t)$ for the t's for which it is finite.
(ii) Using M_X, obtain the quantities: EX, EX^2, and $Var(X)$.
(iii) If the r.v. Y is defined by: $Y = 2 - 3X$, determine $M_Y(t)$ for the t's for which it is finite.

KONTINU

(i)

$$M_X(t) = Ee^{tX} = \int_0^\infty e^{tx} \cdot e^{-x} dx = \int_0^\infty e^{-(1-t)x} dx$$

$$= -\frac{1}{1-t} e^{-(1-t)x} \Big|_0^\infty \quad \text{(provided } t \neq 1\text{)}$$

$$= -\frac{1}{1-t}(0-1) = \frac{1}{1-t} \quad \text{(provided } 1-t > 0 \text{ or } t < 1\text{)}.$$

Thus, $M_X(t) = \frac{1}{1-t}$, $t < 1$.

(ii)

$\frac{d}{dt} M_X(t)|_{t=0} = \frac{d}{dt}(\frac{1}{1-t})|_{t=0} = \frac{1}{(1-t)^2}|_{t=0} = 1 = EX$, $\frac{d^2}{dt^2} M_X(t)|_{t=0} = \frac{d}{dt}(\frac{1}{(1-t)^2})|_{t=0} = \frac{2}{(1-t)^3}|_{t=0} = 2 = EX^2$, so that, by (9), $Var(X) = 2 - 1^2 = 1$.

(iii)

$M_Y(t) = M_{2-3X}(t) = M_{-3X+2}(t) = e^{2t} M_X(-3t) = e^{2t} \frac{1}{1-(-3t)} = \frac{e^{2t}}{1+3t}$, provided $t > -\frac{1}{3}$.