Distribusi Bersama

PDF ↔ *CDF*

PDF ditulis sebagai,

$$\begin{aligned}&f(x_1, x_2, x_3, \ldots, x_n)\\&= P[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n]\end{aligned}$$

Sementara *CDF* terdefinisi sebagai,

$$\begin{aligned}&F(x_1, x_2, x_3, \ldots, x_n)\\&= P[X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n]\end{aligned}$$

CDF diskrit dapat dicari dengan

$$\sum_{x_1}\sum_{x_2}\cdots\sum_{x_n} f(x_1, x_2, \ldots, x_n)$$

Sementara, *PDF* ke *CDF* kontinu,

$$\begin{aligned}&F(x_1, x_2, \ldots, x_n)\\&= \int_{-\infty}^{x_n}\cdots\int_{-\infty}^{x_1} f(t_1, t_2, \ldots, t_n)\, \mathrm{d}t_1 \cdots \mathrm{d}t_n\end{aligned}$$

Sebaliknya, dari *CDF* ke *PDF* kontinu,

$$\begin{aligned}&f(x_1, x_2, \ldots, x_n)\\&= \frac{\partial^k}{\partial x_1 \cdots \partial x_k} F(x_1, x_2, \ldots, x_n)\end{aligned}$$

Peluang Marginal

Untuk dua buah variabel acak (X_1, X_2) dengan *pdf* bersama $f(x_1, x_2)$, maka peluang marginalnya adalah (jika diskrit)

$$f_1(x_1) = \sum_{x_2} f(x_1, x_2)$$
$$f_2(x_2) = \sum_{x_1} f(x_1, x_2)$$

Sementara jika kontinu,

$$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2)\, \mathrm{d}x_2$$
$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2)\, \mathrm{d}x_1$$

Variabel Acak Independen

Variabel dikatakan independen (saling bebas), jika memenuhi

$$\begin{aligned}&f(x_1, x_2, \ldots, x_n)\\&= f(x_1) f(x_2) \cdots f(x_n)\end{aligned}$$

Distribusi Bersyarat

Rumusnya terdefinisi,

$$f(x_1 \mid x_2) = \frac{f(x_1, x_2)}{f_2(x_2)}$$
$$f(x_2 \mid x_1) = \frac{f(x_1, x_2)}{f_1(x_1)}$$
$$f(x_3 \mid x_1, x_2) = \frac{f(x_1, x_2, x_3)}{f(x_1, x_2)}$$

Tinjau pula hubungannya dengan peluang marginal

$$\begin{aligned}f(x_1, x_2) &= f_1(x_1) f(x_2 \mid x_1)\\&= f_2(x_2) f(x_1 \mid x_2)\end{aligned}$$

Jika kedua variabel di atas saling bebas, maka $f(x_2 \mid x_1) = f(x_2)$ dan $f(x_1 \mid x_2) = f(x_1)$.

Sifat Variabel Acak

Sifat Nilai Ekspektasi

Misalkan suatu variabel acak $X = (X_1, X_2, \ldots, X_n)$ memiliki *pdf* gabungan $f(x_1, x_2, \ldots, x_n)$ dan jika $Y = u(X_1, X_2, \ldots, X_n)$ atau sebagai fungsi dari variabel X, maka nilai ekspektasinya adalah

$$\begin{aligned}E(Y) &= E_X[u(X_1, X_2, \ldots, X_n)]\\&= \sum_{x_1}\cdots\sum_{x_n} u(x_1, x_2, \ldots, x_n)\\&\quad\times f(x_1, x_2, \ldots, x_n)\end{aligned}$$

jika X diskrit,

$$\begin{aligned}E(Y) &= E_X[u(X_1, X_2, \ldots, X_n)]\\&= \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} u(x_1, x_2, \ldots, x_n)\\&\quad\times f(x_1, x_2, \ldots, x_n)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n\end{aligned}$$

jika X kontinu.
Hal lain juga perlu diperhatikan,

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

untuk X_1 dan X_2 merupakan variabel acak dengan *pdf* $f(x_1, x_2)$.

Kovarian

Kovarian (atau, *covariant*) ditulis sebagai

$$\mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

atau,

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

Sifat-sifat Kovarian

Adapun sifat-sifat dari kovarian adalah

$$\begin{aligned}\mathrm{Cov}(aX, bY) &= ab\,\mathrm{Cov}(X, Y)\\\mathrm{Cov}(X + a, Y + b) &= \mathrm{Cov}(X, Y)\\\mathrm{Cov}(X, aX + b) &= a\,\mathrm{Var}(X)\end{aligned}$$

Jika dua variabel, X dan Y, saling bebas (*independent*), maka $\mathrm{Cov}(X, Y) = 0$

Varian (Ragam Data) dan Kovarian

Pada dasarnya, untuk dua variabel X dan Y yang memiliki *pdf* $f(x, y)$,

$$\begin{aligned}\mathrm{Var}(X + Y) &= \mathrm{Var}(X) + \mathrm{Var}(Y)\\&\quad + 2\,\mathrm{Cov}(X, Y)\end{aligned}$$

Jika kedua variabel saling bebas (*independent*),

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$$

Untuk banyak variabel X_1, X_2, ..., X_k, kita memiliki

$$\begin{aligned}&\mathrm{Var}\left(\sum_{i=1}^{k} a_i X_i\right)\\&= \sum_{i=1}^{k} a_i^2\,\mathrm{Var}(X_i)\\&\quad + 2\sum\sum_{i<j} a_i a_j\,\mathrm{Cov}(X_i, X_j)\end{aligned}$$

Kalau semua variabel saling independen satu sama lain,

$$\mathrm{Var}\left(\sum_{i=1}^{k} a_i X_i\right) = \sum_{i=1}^{k} a_i^2\,\mathrm{Var}(X_i)$$

Nilai Korelasi

Diberikan dua variabel acak X dan Y dengan ragam σ_X^2 dan σ_Y^2 serta kovarian $\sigma_{XY}^2 = \mathrm{Cov}(X, Y)$, maka koefisien atau nilai korelasinya adalah

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

Nilai Ekspektasi Bersyarat

Jika X dan Y adalah variabel acak yang terdistribusi secara bersama, maka nilai ekspektasi bersyarat Y untuk $X = x$ dinyatakan sebagai

$$E(Y \mid x) = \sum_{y} y f(y \mid x) \tag{1}$$
$$E(Y \mid x) = \int_{-\infty}^{\infty} y f(y \mid x)\, \mathrm{d}y \tag{2}$$

dengan (1) dan (2) untuk kedua variabel yang diskrit dan kontinu, berturut-turut.
Selanjutnya, diketahui pula

$$E[E(Y \mid X)] = E(Y)$$

Untuk X dan Y yang saling bebas (*independent*), $E(Y \mid x) = E(Y)$ dan $E(X \mid y) = E(X)$

Varian (Ragam Data) Bersyarat

Ragam data bersyarat terdefinisi sebagai (menurut aturan nilai ekspektasi),

$$\begin{aligned}\mathrm{Var}(Y \mid x) &= E\left\{\left[Y - E\big(Y \mid x\big)\right]^2 \mid x\right\}\\&= E\left(Y^2 \mid x\right) - [E(Y \mid x)]^2\end{aligned}$$

Jika X dan Y merupakan variabel acak yang terdistribusi secara bersama,

$$\begin{aligned}\mathrm{Var}(Y) &= E_X[\mathrm{Var}(Y \mid X)]\\&\quad + \mathrm{Var}_X[E(Y \mid X)]\end{aligned}$$

Fungsi Pembangkit Momen Gabungan

Jika $X = (X_1, X_2, \ldots, X_n)$, dan jika ada, *MGF* (Fungsi Pembangkit Momen) gabungannya adalah

$$M_X(\mathbf{t}) = E\left[\exp\left(\sum_{k=1}^{n} t_k X_k\right)\right]$$

dengan $\exp(\heartsuit) = e^{\heartsuit}$, e adalah bilangan euler. Serta, $\mathbf{t} = (t_1, t_2, \ldots, t_n)$ dan $-h < t_k < h$ untuk beberapa $h > 0$

Fungsi dari Variabel Acak

Metode CDF

Misal $A_y = \{x \mid u(x) \leq y\}$, maka

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(u(X) \leq y) \\ &= P(X \in A_y) = F_X(A_y) \end{aligned}$$

Misal $X = (X_1, X_2, \ldots, X_k)$ adalah variabel acak kontinu dengan *pdf* bersama $f(x_1, x_2, \ldots, x_k)$, jika $Y = u(X)$ fungsi dari X, maka

$$f_Y(y) = \int \cdots \int_{A_y} f(x_1, \ldots, x_k)\, dx_1 \ldots dx_k$$

Metode Transformasi

one-to-one

Pada kasus diskrit, Misalkan X adalah variabel acak diskrit dengan pdf $f_X(x)$ dan asumsikan bahwa $Y = u(X)$ mendefinisikan transformasi satu-satu. Sehingga misalkan $x = w(y)$ dan $B = \{y \mid f_Y(y) > 0\}$. Maka pdf dari Y adalah

$$f_Y(y) = f_X(w(y)), \quad y \in B$$

Sedangkan untuk kontinu

$$f_Y(y) = f_X(w(y)) \left| \frac{dw(y)}{dy} \right|, \quad y \in B$$

Note: $E_Y(Y) = E_X(u(X))$
Jika transformasi tidak satu-satu, dapat kita partisi A menjadi $A_1, A_2, \ldots, A_n$ sehingga $u(x)$ satu-satu pada A_i. Sehingga untuk kasus diskrit dan kontinu dirumuskan sebagai berikut

$$f_Y(y) = \sum_{i=1}^{n} f_X(w_i(y)), \quad y \in B$$

$$f_Y(y) = \sum_{i=1}^{n} f_X(w_i(y)) \left| \frac{dw_i(y)}{dy} \right|, \quad y \in B$$

Multidimensi

Untuk kasus diskrit sama seperti sebelumnya, hanya perlu menambah peubah acak pada fungsinya. Namun untuk kasus kontinu kita perlu meninjau Jacobian dari fungsi transformasi tersebut.

Misal $X = (X_1, X_2, \ldots, X_n)$ variabel acak kontinu dengan pdf bersama $f_X(x_1, x_2, \ldots, x_n) > 0$ atas A, dan $Y = (Y_1, Y_2, \ldots, Y_n)$ didefinisikan oleh transformasi satu-satu $Y_i = u_i(X_1, X_2, \ldots, X_n)$, maka pdf bersama dari Y adalah

$$f_Y(y_1, y_2, \ldots, y_n) = f_X(x_1, x_2, \ldots, x_n)\, |J|$$

dengan $J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial y_1} & \frac{\partial x_n}{\partial y_2} & \cdots & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$

Jika transformasi tidak satu-satu, dengan cara yang sama yaitu kita partisi A sehingga $u(x)$ satu-satu pada A_i. Kemudian jumlahkan semua pdf nya.

Distribusi Bersama

Tabel Distribusi Diskrit

Nama Distribusi	Notasi dan Parameter	**PDF** Diskrit $f(x)$	**Ekspektasi** $E(X)$	**Varian** $\mathrm{Var}(X)$	**MGF** $M_X(t)$
Bernoulli	$X \sim B(1,p)$	$p^x(1-p)^{1-x}$	p	$p(1-p)$	$1-p+pe^t$
Binomial	$X \sim B(n,p)$	$\binom{n}{x} p^x (1-p)^{n-x}$	np	$np(1-p)$	$(1-p+pe^t)^n$
Negatif Binomial	$X \sim NB(r,p)$	$\binom{x-1}{r-1} p^r (1-p)^x$	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	$\left(\frac{p}{1-(1-p)e^t}\right)^r$
Geometrik	$X \sim G(p)$	$p(1-p)^{x-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{p}{1-(1-p)e^t}$
Hypergeometrik	$X \sim H(n,M,N)$	$\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$	$n\frac{M}{N}$	$n\frac{M}{N}\left(1-\frac{M}{N}\right)\frac{N-n}{N-1}$	-
Poisson	$X \sim P(\mu)$	$\frac{e^{-\mu}\mu^x}{x!}$	μ	μ	$e^{\mu(e^t-1)}$
Uniform Diskrit	$X \sim U(a,b)$	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{tb}-e^{ta}}{t(b-a)}$

Tabel Distribusi Kontinu

Nama Distribusi	Notasi dan Parameter	**PDF** Kontinu $f(x)$	**Ekspektasi** $E(X)$	**Varian** $\mathrm{Var}(X)$	**MGF** $M_X(t)$
Uniform	$X \sim UNIF(a,b)$	$\frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt}-e^{at}}{(b-a)t}$
Normal	$X \sim N(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$	μ	σ^2	$e^{\mu t + \frac{\sigma^2 t^2}{2}}$
Gamma	$X \sim GAM(\theta,\kappa)$	$\frac{1}{\theta^\kappa \Gamma(\kappa)} x^{\kappa-1} e^{-x/\theta}$	$\kappa\theta$	$\kappa\theta^2$	$\left(\frac{1}{1-\theta t}\right)^\kappa$
Exponential	$X \sim EXP(\theta)$	$\frac{1}{\theta} e^{-x/\theta}$	θ	θ^2	$\frac{1}{1-\theta t}$
Weibull	$X \sim WEI(\theta,\beta)$	$\frac{\beta}{\theta^\beta} x^{\beta-1} e^{-(x/\theta)^\beta}$	$\theta\Gamma\left(1+\frac{1}{\beta}\right)$	$\theta^2\left[\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma^2\left(1+\frac{1}{\beta}\right)\right]$	-
Pareto	$X \sim PAR(\theta,\kappa)$	$\frac{\kappa}{\theta(1+x/\theta)^{\kappa+1}}$	$\frac{\theta}{\kappa-1}$	$\frac{\theta^2\kappa}{(\kappa-1)^2(\kappa-2)}$	-