BAB II

TEORI PENUNJANG

Pada bab ini diberikan teori – teori yang menunjang terhadap inti permasalahan diantaranya mengenai statistik terurut dan metode bootstrap.

2.1. Statistik Terurut

Jika $X_1, X_2, ..., X_n$ merupakan sampel acak dari populasi dengan fungsi distribusi komulatif yang kontinu F . Misalkan $X_{1:n}$ merupakan nilai terkecil dari X_i, ... , dan $X_{n:n}$ merupakan nilai terbesar dari X_i, maka $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ disebut statistik terurut. (Dudewicz dan Mishra, 1995)

2.1.1. Fungsi Densitas Bersama Statistik Terurut

Teorema 2.1

Misalkan $X_1, X_2, ..., X_n$ sampel acak berukuran n yang saling bebas dan berdistribusi identik F dengan fungsi densitas f dan $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ menyatakan suatu statistik terurut, maka fungsi densitas bersama dari statistik terurut tersebut adalah :

$$f(x_{1:n}, x_{2:n}, ..., x_{n:n}) = n! \prod_{i=1}^{n} f(x_{i:n}), a < x_{1:n} < x_{2:n} < ... < x_{n:n} < b \qquad 2.1.1$$

$$= 0 \qquad ,\text{lainnya}$$

(Hogg dan Craig)

Bukti:

Bukti:

Misal diambil $n=3$, maka fungsi densitas bersama dari X_1, X_2, X_3 adalah $f(x_1)f(x_2)f(x_3)$. Himpunan A dengan $f(x_1)f(x_2)f(x_3)>0$ adalah gabungan dari himpunan - himpunan kemungkinan :

$$A_1 = \{(x_1, x_2, x_3); a < x_1 < x_2 < x_3 < b\}$$

$$A_2 = \{(x_1, x_2, x_3); a < x_2 < x_1 < x_3 < b\}$$

$$A_3 = \{(x_1, x_2, x_3); a < x_1 < x_3 < x_2 < b\}$$

$$A_4 = \{(x_1, x_2, x_3); a < x_3 < x_1 < x_2 < b\}$$

$$A_5 = \{(x_1, x_2, x_3); a < x_2 < x_3 < x_1 < b\}$$

$$A_6 = \{(x_1, x_2, x_3); a < x_3 < x_2 < x_1 < b\}$$

Jadi himpunan A merupakan gabungan dari $3! = 6$ himpunan kemungkinan.

Misalkan $x_{1:3}$ =minimum dari (x_1, x_2, x_3) , $x_{2:3}$ = nilai tengah (x_1, x_2, x_3), sedangkan $x_{3:3}$ =maksimum dari (x_1, x_2, x_3). Selanjutnya pandang suatu transformasi satu-satu, yang memetakan setiap A_i ke himpunan yang sama yaitu $B = \{(x_{1:3}, x_{2:3}, x_{3:3}); a < x_{1:3} < x_{2:3} < x_{3:3} < b\}$ Maka fungsi invers dari transformasi tersebut :

$$A_1 : x_1 = x_{1:3}, x_2 = x_{2:3}, x_3 = x_{3:3}$$

$$A_2 : x_1 = x_{2:3}, x_2 = x_{1:3}, x_3 = x_{3:3}$$

$$A_3 : x_1 = x_{1:3}, x_2 = x_{3:3}, x_3 = x_{2:3}$$

$$A_4 : x_1 = x_{3:3}, x_2 = x_{1:3}, x_3 = x_{2:3}$$

$$A_5 : x_1 = x_{2:3}, x_2 = x_{3:3}, x_3 = x_{1:3}$$

$$A_6 : x_1 = x_{3:3}, x_2 = x_{2:3}, x_3 = x_{1:3}$$

Sehingga dari fungsi invers tersebut didapatkan suatu transformasi Jacobian . Untuk A_1 transformasi Jacobian yang dihasilkan adalah :

$$J_1 = \begin{vmatrix} \dfrac{\partial x_1}{\partial x_{13}} & \dfrac{\partial x_1}{\partial x_{23}} & \dfrac{\partial x_1}{\partial x_{33}} \\ \dfrac{\partial x_2}{\partial x_{13}} & \dfrac{\partial x_2}{\partial x_{23}} & \dfrac{\partial x_2}{\partial x_{33}} \\ \dfrac{\partial x_3}{\partial x_{13}} & \dfrac{\partial x_3}{\partial x_{23}} & \dfrac{\partial x_3}{\partial x_{33}} \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

Demikian selanjutnya sehingga diperoleh :

$$J_2 = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{vmatrix} = -1$$

$$J_3 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} = -1 \qquad J_4 = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} = 1$$

$$J_5 = \begin{vmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = 1 \qquad J_6 = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = -1$$

Jika $f(x_1, x_2, \ldots, x_n)$ merupakan fungsi densitas bersama dari $X_1, X_2, \ldots, X_n$. Misalkan A ruang berdimensi n dimana $f(x_1, x_2, \ldots, x_n) > 0$, $y_1 = u_1(x_1, x_2, \ldots, x_n), y_2 = u_2(x_1, x_2, \ldots, x_n), \ldots, y_n = u_n(x_1, x_2, \ldots, x_n)$ mendefinisikan transformasi yang memetakan A ke B, dimana A merupakan gabungan dari himpunan – himpunan kemungkinan $A_1, A_2, \ldots, A_k$ sedemikian sehingga

$y_1 = u_1(x_1, x_2, \ldots, x_n), y_2 = u_2(x_1, x_2, \ldots, x_n), \ldots, y_n = u_n(x_1, x_2, \ldots, x_n)$

mendefinisikan transformasi satu – satu yang memetakan tiap A_i ke B.

Kemudian jika J_i menyatakan transformasi Jacobian dari tiap A_i ke B maka fungsi densitas bersama dari $y_1 = u_1(x_1, x_2, \ldots, x_n), \ldots, y_n = u_n(x_1, x_2, \ldots, x_n)$:

$$g(y_1, y_2, \ldots, y_n) = \sum_{i=1}^{k} |J_i| f[u_{1i}(x_1, x_2, \ldots, x_n), \ldots, u_{ni}(x_1, x_2, \ldots, x_n)]$$

Dengan demikian fungsi densitas bersama dari statistik terurut $X_{1:3}, X_{2:3}, X_{3:3}$ adalah:

$$f(x_{1:3}, x_{2:3}, x_{3:3}) = |J_1| f(x_{1:3}) f(x_{2:3}) f(x_{3:3}) + |J_2| f(x_{2:3}) f(x_{1:3}) f(x_{3:3}) + \cdots$$

$$+ |J_6| f(x_{3:3}) f(x_{2:3}) f(x_{1:3})$$

$$= 3! \prod_{i=1}^{3} f(x_{i:3}) \quad , a < x_{1:3} < x_{2:3} < x_{3:3} < b \qquad 2.1.2$$

$$= 0 \qquad , \text{lainnya}$$

Selanjutnya diambil untuk $n = k$, fungsi densitas bersama dari $X_1, X_2, \ldots, X_n$ adalah $f(x_1) f(x_2) \cdots f(x_k)$. Himpunan A dengan $f(x_1) f(x_2) \cdots f(x_k) > 0$ merupakan gabungan dari $k!$ himpunan kemungkinan :

$$A_1 = \{(x_1, x_2, \cdots, x_k); a < x_1 < x_2 < \cdots < x_k < b\}$$

$$\vdots$$

$$A_{k!} = \{(x_1, x_2, \cdots, x_k); a < x_k < x_{k-1} < \cdots < x_1 < b\}$$

Bila $x_{1:k}$: nilai minimum dari x_i ... $x_{k:k}$: nilai maksimum dari x_i. Kemudian pandang suatu transformasi satu - satu yang memetakan setiap A_i ke $B = \{(x_{1:k}, x_{2:k}, \cdots, x_{k:k}); a < x_{1:k} < x_{2:k} < \cdots < x_{k:k} < b\}$. Fungsi invers dari masing – masing titik dalam A_i adalah :

$$A_1 : x_1 = x_{1:k}, x_2 = x_{2:k}, \ldots, x_k = x_{k:k}$$

$$\vdots$$

$$A_{k!} : x_1 = x_{k:k}, x_2 = x_{k-1:k}, \ldots, x_k = x_{1:k}$$

Dari fungsi invers tersebut akan didapatkan transformasi Jacobian yang bernilai ± 1. Sehingga fungsi densitas bersama dari statistik terurut $X_{1:k}, X_{2:k}, \ldots, X_{k:k}$ adalah:

$$f(x_{1:k}, x_{2:k}, \ldots, x_{k:k}) = |J_1| f(x_{1:k}) f(x_{2:k}) \cdots f(x_{k:k}) + \cdots + |J_{k!}| f(x_{k:k}) f(x_{k-1:k}) \cdots f(x_{1:k})$$

$$= k! \prod_{i=1}^{k} f(x_{i:k}) \text{ , } a < x_{1:k} < x_{2:k} < \cdots < x_{k:k} < b \qquad 2.1.3$$

$$= 0 \qquad \text{,lainnya}$$

Untuk n statistik terurut maka fungsi densitas dari $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ adalah :

$$f(x_{1:n}, x_{2:n}, \ldots, x_{n:n}) = n! \prod_{i=1}^{n} f(x_{i:n}) \text{ , } a < x_{1:n} < x_{2:n} < \ldots < x_{n:n} < b$$

$$= 0 \qquad \text{,lainnya}$$

2.1.2 Fungsi Densitas Marginal dari Statistik Terurut

Teorema 2.2

Jika $X_1, X_2, \ldots, X_n$ peubah acak yang saling bebas dan berdistribusi identik F dengan fungsi densitas f dan misalkan $X_{1:n} \le X_{2:n} \le \ldots \le X_{n:n}$ menyatakan suatu statistik terurut, maka fungsi densitas marginal dari $X_{r:n}$ adalah:

$$f_r(x_{r:n}) = \frac{n!}{(r-1)!(n-r)!} [F(x_{r:n})]^{r-1} [1 - F(x_{r:n})]^{n-r} f(x_{r:n}) \text{ , } a < x_{r:n} < b \qquad 2.1.4$$

$$= 0 \qquad \text{,lainnya}$$

(Hogg dan Craig)

Bukti:

Jika diberikan $f(x)$ fungsi densitas yang positif dan kontinu untuk $a < x < b$ dan 0 untuk yang lainnya maka fungsi distribusi $F(x)$ dapat dituliskan dengan :

$$F(x) = 0 \text{ , } x \leq a$$

$$= \int_a^x f(w)dw \text{ , } a < x < b$$

$$= 1 \text{ , } b \leq x$$

Pertama, akan ditunjukkan fungsi densitas marginal dari $X_{n:n}$ yang merupakan nilai maksimum dari $(X_1, X_2, \ldots, X_n)$. Misal diambil $a < x_{n:n} < b$, maka fungsi densitas marginal dari $X_{n:n}$ diberikan dengan :

$$f_n(x_{n:n}) = \int_a^{x_{n:n}} \ldots \int_a^{x_{4:n}} \int_a^{x_{3:n}} \int_a^{x_{2:n}} n! f(x_{1:n}) f(x_{2:n}) \ldots f(x_{n:n}) dx_{1:n} dx_{2:n} \ldots dx_{n-1:n}$$

$$= n! \int_a^{x_{n:n}} \ldots \int_a^{x_{4:n}} \int_a^{x_{3:n}} \left(\int_a^{x_{2:n}} f(x_{1:n}) dx_{1:n} \right) f(x_{2:n}) \ldots f(x_{n:n}) dx_{2:n} \ldots dx_{n-1:n}$$

karena $F(x) = \int_a^x f(w)dw$ maka

$$f_n(x_{n:n}) = n! \int_a^{x_{n:n}} \ldots \int_a^{x_{4:n}} \int_a^{x_{3:n}} F(x_{2:n}) f(x_{2:n}) \ldots f(x_{n:n}) dx_{2:n} \ldots dx_{n-1:n}$$

$$= n! \int_a^{x_{n:n}} \ldots \int_a^{x_{4:n}} \frac{[F(x_{3:n})]^2}{2} f(x_{3:n}) \ldots f(x_{n:n}) dx_{3:n} \ldots dx_{n-1:n}$$

....

$$= n! \frac{[F(x_{n:n})]^{n-1}}{(n-1)} f(x_{n:n})$$

$$= n[F(x_{n:n})]^{n-1} f(x_{n:n}), a < x_{n:n} < b$$

$$= 0 \quad \text{,lainnya}$$

Selanjutnya ditunjukkan fungsi densitas marginal dari $X_{1:n}$. Misal diambil $a < x_{1:n} < b$ maka fungsi densitas marginal dari $X_{1:n}$ adalah:

$$f_1(x_{1:n}) = n! \int_{x_{1:n}}^{b} \cdots \int_{x_{n-3:n}}^{b} \int_{x_{n-2:n}}^{b} \int_{x_{n-1:n}}^{b} f(x_{1:n}) f(x_{2:n}) \ldots f(x_{n:n}) dx_{n:n} dx_{n-1:n} dx_{n-2:n} \ldots dx_{2:n}$$

Bila $F'(x) = f(x)$ dengan $a < x < b$ maka

$$1 - F(x) = F(b) - F(x) = \int_a^b f(w)dw - \int_a^x f(w)dw = \int_x^b f(w)dw$$

sehingga

$$f_1(x_{1:n}) = n! \int_{x_{1:n}}^{b} \cdots \int_{x_{n-3:n}}^{b} \int_{x_{n-2:n}}^{b} f(x_{1:n}) f(x_{2:n}) \ldots f(x_{n-1:n}) [1 - F(x_{n-1:n})] dx_{n-1:n} \ldots dx_{2:n}$$

$$= n! \int_{x_{1:n}}^{b} \cdots \int_{x_{n-3:n}}^{b} f(x_{1:n}) \ldots f(x_{n-2:n}) \frac{[1 - F(x_{n-2:n})]^2}{2} dx_{n-2:n} \ldots dx_{2:n}$$

....

$$= n! \frac{[1 - F(x_{1:n})]^{n-1}}{(n-1)!} f(x_{1:n})$$

$$= n[1 - F(x_{1:n})]^{n-1} f(x_{1:n}), a < x_{1:n} < b$$

$$= 0 \quad \text{, lainnya}$$

Dengan demikian dapat ditunjukkan bahwa fungsi densitas marginal dari statistik terurut ke-r, $X_{r:n}$ adalah:

$$f_r(x_{r:n}) = n! \int_a^{x_{r:n}} \int_a^{x_{r-1:n}} \cdots \int_a^{x_{2:n}} \int_{x_{r:n}}^{b} \int_{x_{r+1:n}}^{b} \cdots \int_{x_{n-1:n}}^{b} f(x_{r:n}) \prod_{\substack{i=1 \\ i \neq r}}^{n} f(x_{i:n}) dx_{n:n} \ldots dx_{r+2:n} dx_{r+1:n} dx_{1:n} \cdots dx_{r-1:n}$$

$$= n! \int_a^{x_{r:n}} \int_a^{x_{r-1:n}} \cdots \int_a^{x_{2:n}} \frac{[1 - F(x_{r:n})]^{n-r}}{(n-r)!} f(x_{r:n}) \prod_{i=1}^{r-1} f(x_{i:n}) dx_{1:n} \ldots dx_{r-2:n} dx_{r-1:n}$$

....

$$= n!\frac{[1-F(x_{r:n})]^{n-r}}{(n-r)!}\frac{[F(x_{r:n})]^{r-1}}{(r-1)!}f(x_{r:n})$$

$$= \frac{n!}{(r-1)!(n-r)!}[F(x_{r:n})]^{r-1}[1-F(x_{r:n})]^{n-r}f(x_{r:n}), a < x_{r:n} < b$$

$$= 0 \quad ,\text{lainnya}$$

Menurut Dudewicz dan Mishra (1995), bila fungsi densitas bersama dari statistik terurut $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ diintegralkan terhadap semua peubah kecuali $X_{r:n}$ dan $X_{s:n}$ $(r < s)$ akan diperoleh :

Akibat 2.1.1

Fungsi densitas marginal dari dua statistik terurut, misalkan $X_{r:n}$ dan $X_{s:n}$ $(r < s)$ adalah:

$$f_{rs}(x_{r:n}, x_{s:n}) = \frac{n!}{(n-s)!(r-1)!(s-r-1)!}[F(x_{r:n})]^{r-1}[F(x_{s:n})-F(x_{r:n})]^{s-r-1}[1-F(x_{s:n})]^{n-s}$$
$$\times f(x_{r:n})f(x_{s:n}), a < x_{r:n} < x_{s:n} < b \qquad 2.15$$

$$= 0 \qquad ,\text{lainnya}$$

2.1.3 Nilai Harapan (Ekspektasi) dari Statistik Terurut

Teorema 2.3

Jika $X_{1:n} \le X_{2:n} \le \ldots \le X_{n:n}$ menyatakan suatu statistik terurut, maka nilai ekspektasi dari $X_{r:n}$ adalah :

$$\mu_{r:n} = r\binom{n}{r}\int_0^1 F^{-1}(u)u^{r-1}(1-u)^{n-r}du \qquad 2.1.6$$

(David, 1981)

Bukti :

$$\mu_{r:n} = E(X_{r:n})$$

$$= \int_{-\infty}^{\infty} x_{r:n} f_r(x_{r:n}) dx_{r:n}$$

$$= \int_{-\infty}^{\infty} x_{r:n} \frac{n!}{(n-r)!(r-1)!} [F(x_{r:n})]^{r-1} [1 - F(x_{r:n})]^{n-r} f(x_{r:n}) dx_{r:n}$$

$$= \int_{-\infty}^{\infty} x_{r:n} \frac{n!}{(n-r)!(r-1)!} \frac{r}{r} [F(x_{r:n})]^{r-1} [1 - F(x_{r:n})]^{n-r} dF(x_{r:n})$$

Bila $0 \le F(x_{r:n}) \le 1$ dan berdasar sifat transformasi integral probabilitas yaitu bila X merupakan variabel random yang mempunyai fungsi distribusi F yang kontinu, maka variabel random Y yang dihasilkan dari transformasi $Y = F(x)$ mempunyai distribusi probabilitas uniform pada $(0,1)$. Sehingga $u = F(x_{r:n})$ merupakan sampel random dari populasi uniform pada $(0,1)$. Selanjutnya jika U merupakan variabel random dari populasi uniform pada $(0,1)$ dan $F(X_{r:n})$ mempunyai invers $F^{-1}(.)$, maka $X_{r:n} = F^{-1}(U)$. Sehingga

$$\mu_{r:n} = \int_{0}^{1} r\binom{n}{r} F^{-1}[F(x_{r:n})] \, [F(x_{r:n})]^{r-1} [1 - F(x_{r:n})]^{n-r} dF(x_{r:n})$$

$$= r\binom{n}{r} \int_{0}^{1} F^{-1}(u) u^{r-1} (1-u)^{n-r} du$$

2.1.4. Variansi Statistik Terurut

Teorema 2.4

Jika $X_1, X_2, ..., X_n$ sampel acak yang saling bebas dan berdistribusi identik F dengan fungsi densitas f dan misalkan $X_{1:n} \le X_{2:n} \le ... \le X_{n:n}$ menyatakan suatu statistik terurut maka variansi dari $X_{r:n}$ adalah :

$$\sigma_{r:n}^2 = \int_0^1 r\binom{n}{r}\left[F^{-1}(u) - \mu_{r:n}\right]^2 u^{r-1}(1-u)^{n-r}du \qquad 2.1.7$$

(David, 1981)

Bukti :

$$\sigma_{r:n}^2 = E(X_{r:n} - \mu_{r:n})^2$$

$$= \int_{-\infty}^{\infty}(x_{r:n} - \mu_{r:n})^2 f_r(x_{r:n})dx_{r:n}$$

$$= \int_{-\infty}^{\infty}(x_{r:n} - \mu_{r:n})^2 \frac{n!}{(n-r)!(r-1)!}[F(x_{r:n})]^{r-1}[1-F(x_{r:n})]^{n-r} f(x_{r:n})dx_{r:n}$$

$$= \int_{-\infty}^{\infty}(x_{r:n} - \mu_{r:n})^2 \frac{n!}{(n-r)!(r-1)!}\frac{r}{r}[F(x_{r:n})]^{r-1}[1-F(x_{r:n})]^{n-r} f(x_{r:n})dx_{r:n}$$

Bila $0 \le F(x_{r:n}) \le 1$ dan $u = F(x_{r:n})$ maka,

$$\sigma_{r:n}^2 = r\binom{n}{r}\int_0^1 \left[F^{-1}(u) - \mu_{r:n}\right]^2 u^{r-1}(1-u)^{n-r}du$$

2.1.5. Kovariansi Statistik Terurut

Teorema 2.5

Jika $X_1, X_2, ..., X_n$ sampel acak yang saling bebas dan berdistribusi identik F dengan fungsi densitas f dan misalkan $X_{1:n} \le X_{2:n} \le ... \le X_{n:n}$ menyatakan suatu statistik terurut maka kovariansi dari $X_{r:n}$ dan $X_{s:n}$ adalah:

$$\sigma_{rs:n} = \int_0^1\int_0^{u_s}\left(F^{-1}(u_r) - \mu_{r:n}\right)\left(F^{-1}(u_s) - \mu_{s:n}\right)_n c_{rs}[u_r]^{r-1}[u_s - u_r]^{s-r-1}[1-u_s]^{n-s}du_r du_s$$

(David, 1981)

Bukti :

$$\sigma_{rs:n} = E\{(X_{r:n} - \mu_{r:n})(X_{s:n} - \mu_{s:n})\}$$

$$= \int_{-\infty}^{\infty}\int_{\infty}^{x_{s:n}} (x_{r:n} - \mu_{r:n})(x_{s:n} - \mu_{s:n}) f_{rs}(x_{r:n}, x_{s:n}) dx_{r:n} dx_{s:n}$$

$$= \int_{-\infty}^{\infty}\int_{\infty}^{x_{s:n}} (x_{r:n} - \mu_{r:n})(x_{s:n} - \mu_{s:n})_n c_{rs} [F(x_{r:n})]^{r-1} [F(x_{s:n}) - F(x_{r:n})]^{s-r-1}$$
$$\times [1 - F(x_{s:n})]^{n-s} f(x_{r:n}) f(x_{s:n}) dx_{r:n} dx_{s:n}$$

dengan $-\infty < x_{r:n} < x_{s:n} < \infty$

bila $0 \le F(x_{r:n}) \le F(x_{s:n}) \le 1$, $u_r = F(x_{r:n})$, dan $u_s = F(x_{s:n})$ maka persamaan di atas menjadi :

$$= \int_0^1 \int_0^{u_s} (F^{-1}(u_r) - \mu_{r:n})(F^{-1}(u_s) - \mu_{s:n})_n c_{rs} [u_r]^{r-1} [u_s - u_r]^{s-r-1} [1 - u_s]^{n-s} du_r du_s$$

2.2 Metode Bootstrap

Metode bootstrap pertamakali diperkenalkan oleh B. Efron pada tahun 1979. Metode ini berfungsi untuk mengestimasi berbagai kuantitas statistik, menentukan interval konfidensi, dan untuk menentukan fungsi densitas non parametrik. Pendekatan bootstrap dasar adalah memperlakukan sampel sebagai populasi dan menerapkan sampling untuk membangkitkan dugaan empiris bagi distribusi sampling statistik.

Pandang $X_1, X_2, \ldots, X_n$ sebagai sampel acak berukuran n dari populasi dengan fungsi distribusi kontinu F yang tidak diketahui, dan misalkan

$$\theta = \theta(F) \qquad 2.2.1$$

merupakan parameter yang akan diestimasi. Parameter θ dapat dipandang sebagai nilai fungsi sebenarnya dari F. Estimator dari θ yaitu $\hat{\theta}$ merupakan suatu fungsi dari $X_1, X_2, ..., X_n$. Distribusi F yang tidak diketahui dapat diestimasi menggunakan fungsi distribusi empiris $\hat{F}_n$ berdasar sampel $X_1, X_2, ..., X_n$ dengan peluang $1/n$ untuk setiap X_i. Fungsi distribusi empiris $\hat{F}_n$ didefinisikan sebagai :

$$\hat{F}_n(x) = \frac{1}{n}\{\text{ banyaknya } X_i \leq x\},\ 1 \leq i \leq n, \text{ untuk } -\infty < x < \infty.$$

Sampel bootstrap $X^* = (X_1^*, X_2^*, ..., X_n^*)$ didefinisikan sebagai sampel random berukuran n yang diambil dari $\hat{F}_n$ dengan pengembalian (B. Efron, 1993). Jadi bila dipunyai sampel random berukuran n akan didapat kemungkinan sampel bootstrap sebanyak n^n, dari tiap sampel bootstrap ini dapat ditentukan $\hat{\theta}^*$. Seluruh kemungkinan sampel bootstrap ini dinamakan dengan jumlah sampel bootstrap ideal.

Perhitungan $\hat{\theta}^*$ berdasar semua kemungkinan sampel bootstrap memerlukan waktu yang cukup lama. Sehingga untuk mencapai efisiensi dalam perhitungan digunakan suatu metode pendekatan yaitu simulasi Monte Carlo, dengan simulasi tersebut prosedur resampling pada metode bootstrap ideal dapat dikurangi menjadi $n \leq B \leq n^n$, sejumlah B yang cukup besar tetapi jauh lebih kecil jika dibandingkan dengan jumlah sampel bootstrap ideal. Menurut Helmers dan Putter besarnya nilai B biasanya berkisar antara 100 sampai 1000 , dan pengambilan nilai B tergantung dari inferensi yang diambil. Misalnya untuk estimasi variansi besarnya nilai B adalah 200.

Langkah – langkah dasar metode bootstrap menurut Efron :

a. Menentukan distribusi empiris, $\hat{F}_n(x)$ bagi sampel dengan peluang 1/n untuk masing – masing X_i

b. Menentukan sampel bootstrap $X_1^*, X_2^*, ..., X_n^*$ yang diambil dari X_i dengan pengembalian

c. Menentukan replikasi bootstrap $\hat{\theta}*$, berdasar sampel bootstrap

d. Ulangi langkah b dan c sebanyak B kali

e. Berikan distribusi probabilitas bagi $B\hat{\theta}*$ dengan menempatkan probabilitas $1/B$ bagi masing – masing $\hat{\theta}_1*, \hat{\theta}_2*, ..., \hat{\theta}_B*$. Distribusi ini adalah estimasi bootstrap untuk distribusi sampling $\hat{\theta}$.