Sampling

مجموع‌های تصادفی (Random Sums)

$$\boxed{Y = \sum_{i=1}^{N} X_i}$$

$$E[Y] = E\big[E[Y|N]\big] = E\Big[E\Big[\sum_{i=1}^{N} X_i \,\Big|\, N\Big]\Big] = E\Big[\sum_{i=1}^{N} E[X_i]\Big]$$

$$= E\big[N\,E[X]\big] = E[X]\,E[N]$$

میانگین مجموع تصادفی

واریانس مجموع تصادفی

$$E[y^2] = E\left[E[y^2|N]\right] = E\left[E\left[\left(\sum_{i=1}^{N} x_i\right)^2 \Big| N\right]\right]$$

$$= E\left[E\left[\sum_{i=1}^{N}\sum_{j=1}^{N} x_i x_j \Big| N\right]\right] = E\left[\sum_{i=1}^{N}\sum_{j=1}^{N} E[x_i x_j]\right]$$

$$= E_N\left[N\, E[x^2] + (N^2 - N)\, E^2[x]\right]$$

$$\sigma_x^2 = E[x^2] - E^2[x]$$

$$= \mu_N E[x^2] + E^2[x]\left(E[N^2] - \mu_N\right)$$

$$= \mu_N\left(\sigma_x^2 + \mu_x^2\right) + \mu_x^2\left(\sigma_N^2 + \mu_N^2 - \mu_N\right)$$

$$\mathrm{var}(y) = E[y^2] - E^2[y] = E[y^2] - (\mu_x \mu_N)^2$$

$$Y = \sum_{i=1}^{N} X_i \longrightarrow \text{Random sum}$$

X_i i.i.d.

$$E[Y] = \underline{E[N]\,E[X]}$$

$$E[Y] = E\left[E[Y|N]\right]$$

$\mathrm{var}(Y)$

مثال

- فرستنده‌های A و B پیام‌هایی را با توزیع‌های پواسون با نرخ‌های λ_A و λ_B به یک گیرنده خاص می‌فرستند. تعداد کلمات هر پیام مستقل از یکدیگر و با تابع جرمی احتمال

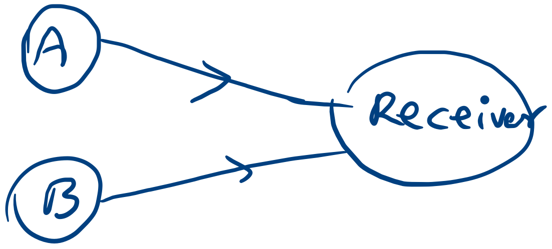
$$P_W(w=1)=\frac{1}{3}\ , \qquad P_W(w=2)=\frac{1}{2}, \qquad P_W(w=3)=\frac{1}{6}$$

توزیع شده است.

الف) احتمال این که در بازه‌ای به طول t دقیقاً ۹ پیام به گیرنده برسد، چقدر است؟

$X \sim Poi(\lambda_A)$

$Y \sim Poi(\lambda_B)$



$Z = X+Y \Rightarrow Z \sim Poi(\lambda_A+\lambda_B)$

$$P(Z=9) = \frac{((\lambda_A+\lambda_B)t)^9}{9!} e^{-(\lambda_A+\lambda_B)t}$$

ادامه مثال

- ب) امید ریاضی تعداد کلمات دریافتی N در بازه‌ای به طول t را محاسبه کنید.

$$N = W_1 + W_2 + \cdots + W_Z = \sum_{i=1}^{Z} W_i$$

$$\boxed{E[N] = E[Z]\,E[W]} = (\lambda_A + \lambda_B)t\left(\frac{1}{3}\times 1 + \frac{1}{2}\times 2 + \frac{1}{6}\times 3\right)$$

نمونه‌برداری (Sampling)

- به دنباله متغیرهای تصادفی $i.i.d.$ $(X_1, X_2, \dots, X_n)$ که از یک جامعه (population) آماری با توزیع F انتخاب شده باشند، یک نمونه (sample) از توزیع F می‌گوییم.

  - چون X_iها مستقل هستند، داریم:

$$f_{X_i}(x) = f_X(x)$$

$$\Rightarrow E(X_i) = E(X) = \mu \quad : \quad \text{میانگین جامعه}$$

$$\Rightarrow \operatorname{var}(X_i) = \operatorname{var}(X) = \sigma^2 \quad : \quad \text{واریانس جامعه}$$

$f_X(x)$

$\{X_1, X_2, \ldots, X_n\} \longrightarrow$ Sample

i.i.d.

نمونه

Population $f_X(x)$

$\{X_1, \ldots, X_n\} \longrightarrow$ sample

P

$1-P$

میانگین نمونه (Sample Mean)

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \{X_1, \dots, X_n\}$$

R.V.

Sample size

$$E[\bar{X}] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = E[X] = \mu$$

$$\mathrm{var}(\bar{X}) = \mathrm{var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\, n\, \mathrm{var}(X) = \frac{\sigma^2}{n}$$

$$\lim_{n\to\infty} \bar{X} = \mu$$

تخمین‌گر ناریب (Unbiased Estimator)

$\theta \qquad \hat{\theta}$

$E[\hat{\theta}] = \theta \longrightarrow$ $\hat{\theta}$ is an unbiased estimator for θ

$E[\bar{X}] = \mu$

$E[\hat{\theta}] = \theta + 2 \qquad E[\hat{\theta}] = 2\theta$

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

μ

$$E[\bar{X}] = \mu$$

P

$$\boxed{\operatorname{var}(\bar{X}) = \frac{\sigma^2}{n}}$$

واریانس نمونه (Sample Variance)

$\{X_1, \dots, X_n\}$ $\quad f_X(x)$

$\sigma_X^2 = ?$

unbiased

$$\boxed{S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

biased

$$\boxed{S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2}$$

$$E[S^2] \neq \sigma^2$$

$$E[S^2] = \frac{1}{n}\sum_{i=1}^{n}\underbrace{E[(X_i - \mu)^2]}_{var(X_i)} = \frac{1}{n} \times n \, Var(X) = \sigma_X^2$$

$$\boxed{S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$x_1, \dots, x_n, \bar{x}$

$$\boxed{(n-1)\,E[S^2]} = E\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right] = E\left[\sum_{i=1}^{n}(x_i^2+\bar{x}^2-2\bar{x}x_i)\right]$$

$$= \sum_{i=1}^{n}\Big(\underbrace{E[x_i^2]}_{\sigma^2+\mu^2} + \underbrace{E[\bar{x}^2]}_{\frac{\sigma^2}{n}+\mu^2} - 2\,E[\bar{x}x_i]\Big)$$

$$\sigma_x^2 = \underbrace{E[x^2]} - E^2[x]$$

$$= n(\sigma^2+\mu^2) + n\left(\frac{\sigma^2}{n}+\mu^2\right) - \underbrace{2\,E\Big[\bar{x}\underbrace{\sum_{i=1}^{n}x_i}_{n\bar{x}}\Big]}_{2n\,\underbrace{E[\bar{x}^2]}_{\frac{\sigma^2}{n}+\mu^2}}$$

$$= n(\sigma^2+\cancel{\mu^2}) - n\left(\frac{\sigma^2}{n}+\cancel{\mu^2}\right) = (n-1)\,\sigma^2$$

$$\Longrightarrow \boxed{E[S^2] = \sigma^2}$$

اثبات نااریب بودن واریانس نمونه

آماره رتبه (Order Statistic)

statistic

$X_1, \ldots, X_n$

$X_{(1)}, \ldots, X_{(n)}$

$Y_1, \ldots, Y_n$

آماره رتبه اول

$f_{Y_1}(y) = ?$

$$F_{Y_1}(y) = P(Y_1 \le y) = 1 - P(Y_1 > y) = 1 - P(\underbrace{X_1 > y}, \dots, \underbrace{X_n > y})$$

$$= 1 - \underbrace{P(X_1 > y)} \cdots P(X_n > y)$$

$$= 1 - (1 - F_X(y))^n$$

$$\boxed{f_{Y_1}(y) = n(1 - F_X(y))^{n-1} \, f_X(y)}$$

آماره رتبه آخر

$$f_{Y_n}(y) = ?$$

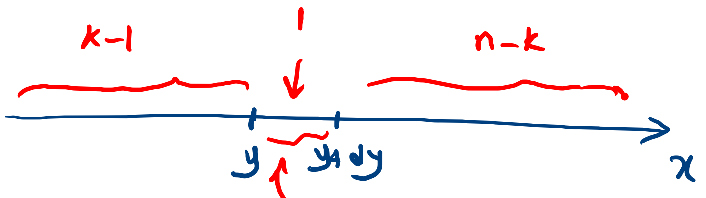
$$F_{Y_n}(y) = P(Y_n < y) = P(X_1 \leq y, \dots, X_n \leq y) = (F_X(y))^n$$

$$\boxed{f_{Y_n}(y) = n(F_X(y))^{n-1} f_X(y)}$$

آماره رتبه k

$$f_{Y_k}(y) = ?$$

$$f_{Y_k}(y)\,\cancel{dy} = P\overbrace{(y \le Y_k \le y + dy)}$$



$$= \binom{n}{k-1,\,1,\,n-k} \big(P(X \le y)\big)^{k-1} \boxed{f_X(y)\,dy} \big(P(X \ge y+dy)\big)^{n-k}$$

$$= \frac{n!}{(k-1)!\,1!\,(n-k)!} \big(F_X(y)\big)^{k-1} f_X(y)\,\cancel{dy} \big(1 - F_X(y+dy)\big)^{n-k}$$

$$\boxed{f_{Y_k}(y) = \frac{n!}{(k-1)!\,(n-k)!} \big(F_X(y)\big)^{k-1} f_X(y) \big(1 - F_X(y)\big)^{n-k}}$$