Stats 426 Lecture 7

Thomas Cohn

2/10/21

Expected Value of Functions of Random Variables

Suppose $X_1, \dots, X_n$ are random variables, and $Y = g(X_1, \dots, X_n)$. Then

- Discrete case: For pmf $p(x_1, \dots, x_n)$, we have $E(Y) = \sum_{x_i \in X_i} g(x_1, \dots, x_n) p(x_1, \dots, x_n)$, if $E(|g(X_1, \dots, X_n)|)$ is finite.
- Continuous case: For pdf $f(x_1, \dots, x_n)$, we have $E(Y) = \int_{X_i} g(x_1, \dots, x_n) p(x_1, \dots, x_n)$, if $E(|g(X_1, \dots, X_n)|)$ is finite.

Ex: If $E(X_i) = \mu$, for $i = 1, \dots, n$, then $E(X_1 + \cdots + X_n) = n\mu$, and $E(\bar{X}) = E((X_1 + \cdots + X_n)/n) = \mu$.

If X, Y are independent random variables, and g, h functions, then $E(g(X)h(X)) = E(g(X))E(h(X))$; $E(XY) = E(X)E(Y)$.

Ex: Suppose $X_1 \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathrm{N}(\mu_2, \sigma_2^2)$, with $X_1 \perp X_2$. Then $E(X_1 X_2) = E(X_1)E(X_2) = \mu_1 \mu_2$.

Variance and Standard Deviation

Defn: Let X be a random variable with mean μ. Then the variance is

$$\mathrm{Var}(X) = E[(X - \mu)^2]$$

We denote the standard deviation of X as σ_X, where $\mathrm{Var}(X) = \sigma_X^2$, i.e., $\sqrt{\mathrm{Var}(X)} = \sigma_X$.

Properties:

- $\mathrm{Var}(X) = E[(X - \mu)^2] = E(X^2 - 2\mu X + \mu^2) = E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2$
- $\mathrm{Var}(a + bX) = E((a + bX - a - b\mu)^2) = E(b^2(X - \mu)^2) = b^2 \mathrm{Var}(X)$.

Exer: Derive the following:

- $X \sim \mathrm{Ber}(p) \rightsquigarrow \mathrm{Var}(X) = p(1-p)$
- $X \sim \mathrm{Bin}(n, p) \rightsquigarrow \mathrm{Var}(X) = np(1-p)$
- $X \sim \mathrm{Geom}(p) \rightsquigarrow \mathrm{Var}(X) = (1-p)/p^2$
- $X \sim \mathrm{Poisson}(\lambda) \rightsquigarrow \mathrm{Var}(X) = \lambda$
- $X \sim \mathrm{Unif}(a, b) \rightsquigarrow \mathrm{Var}(X) = (b-a)^2/12$
- $X \sim \mathrm{Exp}(\lambda) \rightsquigarrow \mathrm{Var}(X) = 1/\lambda^2$
- $X \sim \mathrm{Gamma}(\lambda, \alpha) \rightsquigarrow \mathrm{Var}(X) = \alpha/\lambda^2$
- $X \sim \mathrm{Beta}(a, b) \rightsquigarrow \mathrm{Var}(X) = \frac{ab}{(a+b)^2(a+b+1)}$
- $X \sim \mathrm{N}(\mu, \sigma^2) \rightsquigarrow \mathrm{Var}(X) = \sigma^2$

The standard deviation measures how "spread out" the distribution is about the mean. For the normal distribution, $\mu \pm 2\sigma$ and $\mu \pm 3\sigma$ covers 95.44% and 99.74% of the mass of the distribution.

Thm: (Chebychev Inequality) Say X is a random variable, with $E(X) = \mu$ and $|Var(X) = \sigma^2$.
For $k > 0$, $P(|X - \mu| \geq k) \leq \sigma^2/k^2$, or equivalently, $P(|X - \mu| \geq k\sigma) \leq 1/k^2$.

Proof: Let $R + \{x : |x - \mu| \geq k\}$. Then

$$P(|X - \mu| \geq k) = \int_R 1 \cdot f(x)\,dx \leq \int_R \frac{(x-\mu)^2}{k^2} f(x)\,dx = \frac{1}{k^2} \int_R (x-\mu)^2 f(x)\,dx = \frac{\sigma^2}{k^2}$$

$\square$