Intro Prob Lecture Notes

William Sun

March 6, 2017

Law of the Unconscious Statistician

- Note: For discrete random variables, see Proposition 4.1 of the textbook
- If X is discrete and $G : \mathbb{R} \to \mathbb{R}$, then
  - $\mathbb{E}(\mathbf{g}(\mathbf{X})) = \sum_{\mathbf{x}} \mathbf{g}(\mathbf{x})\mathbf{P}(\mathbf{X} = \mathbf{x})$ when the expectation exists
- Some consequences:
  - Take $g(x) = ax + b$ where $a, b \in \mathbb{R}$ are fixed

$$\begin{aligned}
\mathbb{E}(aX + b) &= \sum_x (ax + b)P(X = x) \\
&= \sum_x (axP(X = x) + bP(X = x)) \\
&= \sum_x axP(X = x) + \sum_x bP(X = x) \\
&= a\sum_x xP(X = x) + b\sum_x P(X = x) \\
&= a\mathbb{E}(X) + b * 1
\end{aligned}$$

  - Linearity of Expectation #1
    - $\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$
  - Linearity of Expectation #2
    - $\mathbb{E}(X_1 + X_2 + \cdots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \cdots + \mathbb{E}X_n)$
    - Expectation of a sum is the sum of the individual expected values

    - For any random variables for which $\mathbb{E}(X_i)$ exists for all i
    - *To be proved later*

- $\mathbb{E}(X) = \mu$, the mean of X
  - Allows us to define the variance

Variance

- The variance of an random variable: $G(x) = (x - \mu)^2$, where $\mu = \mathbb{E}(X)$
  - So $\mathbb{E}(\{X - \mu\}^2) := Var(X)$
  - E.g.: $p(-2) = .5, p(0) = .25, p(4) = .25$
  - $\mu = -2(.50) + 0(.25) + 4(.25) = 0$
    - "Center of mass"

$$
\begin{aligned}
Var(X) &= \mathbb{E}(\{X - \mu\}^2) \\
&= \sum_x (x - \mu)^2 P(X = x) \\
&= (-2 - 0)^2 * .5 + (0 - 0)^2 * .25 + (4 - 0)^2 * .25 \\
&= 2 + 0 + 4 \\
&= 6
\end{aligned}
$$

- A form of the $Var(X)$ more amenable to calculations: $\mathbf{Var(X) = \mathbb{E}(X^2) - \mu^2}$
- To see that this formula is equivalent to our definition:

$$
\begin{aligned}
\mathbb{E}[(X-\mu)^2] &= \sum_x (x-\mu)^2 P(X=x) \\
&= \sum_x (x^2 - 2\mu x + \mu^2) P(X=x) \\
&= \sum_x x^2 P(X=x) - 2\mu \sum_x x P(X=x) + \mu^2 \sum_x P(X=x) \\
&= \mathbb{E}(X^2) - 2\mu^2 + \mu^2 \\
&= \mathbb{E}(X^2) - \mu^2 \\
&= \mathbb{E}(X^2) - \{\mathbb{E}(X)\}^2 \geq \mathbf{0}
\end{aligned}
$$

- Lyapunov's Inequality: $\{\mathbb{E}(X)\}^2 \leq \mathbb{E}(X^2)$
- Just knowing the mean and variance of a distribution tells us a lot about the probability mass function

Last time: $X \sim$ geometric(p) $\implies \mu = \frac{1}{p} = \mathbb{E}(X)$

- Let's first find $\mathbb{E}(X^2)$:

$$
\begin{aligned}
\mathbb{E}(X^2) &= \sum_{x=1}^{\infty} x^2 p(1-p)^{x-1} \\
\mathbb{E}(X^2) &= p + 4p(1-p) + 9p(1-p)^2 + \dots \\
(1-p)\mathbb{E}(X^2) &= p(1-p) + 4p(1-p)^2 + \dots \\
p\mathbb{E}(X^2) &= p + 3p(1-p)^2 + 5p(1-p)^2 + 7p(1-p)^3 + \dots \\
\mathbb{E}(X^2) &= 1 + 3(1-p) + 5(1-p)^2 + 7(1-p)^3 + \dots \\
(1-p)\mathbb{E}(X^2) &= (1-p) + 3(1-p)^2 + 5(1-p)^3 + \dots \\
p\mathbb{E}(X^2) &= 1 + \{2(1-p) + 2(1-p)^2 + 2(1-p)^3 \dots\} \\
&= \text{(Term in curly braces is a geometric series)} \\
p\mathbb{E}(X^2) &= 1 + \frac{2(1-p)}{p} = \frac{2-p}{p}
\end{aligned}
$$

$$\sigma^{\mathbf{2}} = \mathbf{Var(X)} = \frac{\mathbf{2-p}}{\mathbf{p^2}} - \{\frac{\mathbf{1}}{\mathbf{p}}^{\mathbf{2}}\} = \frac{\mathbf{1-p}}{\mathbf{p^2}}$$

- Notation: σ: standard deviation of random variable X
  - Therefore, the square root of the variance is the standard deviation

Chebyshev Rule

- If we know the mean and variance of a distribution...
- Let $k \geq 1$

$$(|X - \mu| \leq k\sigma)$$

  - Lets us know the number of values that are within k standard deviations of μ
  - $\mu - 2\sigma \leq X \leq \mu + 2\sigma$ since $-2\sigma \leq X - \mu \leq 2\sigma$
- The amount of mass outside k standard deviations is $\frac{\mathbf{1}}{\mathbf{k^2}}$
- With $k = 2$, at least 75% of the mass has to be within 2 standard deviations
- $k = 3$ implies $\geq \frac{8}{9}$
- $k = 5$ implies $\geq 96\%$