

Probability Theory: Lecture #13

① What are the expectation and variances for some of our familiar r.v.? Well, we can often exploit the properties with respect to addition to make the calculations simpler.

ex: Suppose $\bar{X} \sim \text{Bernoulli}(p)$. Then $i.e. \bar{X} = \begin{cases} 1, & \text{prob} = p \\ 0, & \text{prob} = 1-p \end{cases}$

$$\mathbb{E}(\bar{X}) = \sum_{\omega} \bar{X}(\omega) P(\omega)$$

$$= 1 \cdot p + 0 \cdot (1-p)$$

$$= p.$$

$$\text{Var}(\bar{X}) = \sum_{\omega} [\bar{X}(\omega) - \mathbb{E}(\bar{X})]^2 \cdot P(\omega)$$

$$= [1-p]^2 \cdot p + p^2 \cdot (1-p)$$

$$= p[1-p][1-p+p]$$

$$= p(1-p).$$

From this, we can immediately go to a binomial r.v.:

ex: Suppose $\bar{X} \sim \text{Bin}(n, p)$. Then we could directly compute \mathbb{E} , Var from the definition, noting that $P(\bar{X}=k) = \binom{n}{k} p^k (1-p)^{n-k}$.

②

This gets tedious, and we can note that $\bar{X} = \sum_{i=1}^n \bar{Y}_i$, where each \bar{Y}_i is an iid Bernoulli(p) r.v. So, we get

$$\mathbb{E}(\bar{X}) = \mathbb{E}\left(\sum_{i=1}^n \bar{Y}_i\right)$$

$$= \sum_{i=1}^n \mathbb{E}(\bar{Y}_i)$$

$$= \sum_{i=1}^n p$$

$$= np.$$

Note that we didn't need independence of the $\{\bar{Y}_i\}_{i=1}^n$ here. For variance, we do need it:

$$\text{Var}(\bar{X}) = \text{Var}\left(\sum_{i=1}^n \bar{Y}_i\right)$$

assumption
of
independence \downarrow

$$= \sum_{i=1}^n \text{Var}(\bar{Y}_i)$$

$$= n \cdot p(1-p).$$

Ex: Suppose $\bar{X} \sim \text{Poisson}(\lambda)$. Then $\mathbb{E}(\bar{X}) = \sum_{k=0}^{\infty} k \cdot P(\bar{X}=k)$

$$= \sum_{k=0}^{\infty} k \cdot \frac{e^{-\lambda} \cdot \lambda^k}{k!}$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} k \frac{\lambda^k}{k!}$$

Note that $\sum_{k=0}^{\infty} k \cdot \frac{\lambda^k}{k!} = [\lambda e^{\lambda x}] \Big|_{\lambda=0}$, where we interpret $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$ as

the Taylor expansion of $e^{\lambda x}$ about 0, evaluated at 0:

$$e^{\lambda x} = \sum_{k=0}^{\infty} \frac{\lambda^k \cdot x^k}{k!}$$

$$\Rightarrow \frac{d}{dx} [e^{\lambda x}] = \left[\sum_{k=1}^{\infty} \frac{\lambda^k \cdot k \cdot x^{k-1}}{k!} \right]$$

$$\lambda e^{\lambda x} = \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k}{k!} \cdot x^{k-1}$$

Plugging in $x=1$ gives $\lambda e^{\lambda} = \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k}{k!}$

So, we conclude $\mathbb{E}(X) = e^{-\lambda} \cdot \lambda e^{\lambda} = \lambda$.

This could have been seen in a different (perhaps shorter) way by noting that $\text{Poisson}(\lambda)$ is the limit of a sequence of $\bar{X}_n \sim \text{Bin}(n, \frac{\lambda}{n})$, each \bar{X}_n satisfies $\mathbb{E}(\bar{X}_n) = n \cdot \frac{\lambda}{n} = \lambda$, so therefore the limit is

$n \rightarrow \infty$ (namely $\text{Poisson}(\lambda)$) must also have expected value λ . (4)

//

• Just as we considered notion of conditional probability, we may consider notion of conditional expectation. Just swap $P(X=w)$ with $P(X=w| \dots)$ in the definition.

Defn.: Let X be ~~a.s.m.~~ a r.v. with sample space $\Omega = \{x_i\}_{i=1}^{\infty}$, and let E be any event. Then $\mathbb{E}(X|E) = \sum_{i=1}^{\infty} x_i \cdot P(X=x_i|E)$.

• A partition of Ω into events $\{F_j\}_{j=1}^{\infty}$ allows one to compute $\mathbb{E}(X)$ by conditioning with respect to the F_j , then summing:

Theorem (E and total probability): Let X be ~~a.s.m.~~ a random variable on $\Omega = \{x_i\}_{i=1}^{\infty}$ a discrete space. Let $\{F_j\}_{j=1}^{\infty}$ be a partition of Ω . Then

$$\mathbb{E}(X) = \sum_{j=1}^{\infty} \mathbb{E}(X|F_j) \cdot P(F_j).$$

Prof.: By definition, the RHS equals $\sum_{j=1}^{\infty} \mathbb{E}(X|F_j) \cdot P(F_j)$

$$= \sum_{j=1}^{\infty} \left(\sum_{i=1}^{\infty} x_i \cdot P(X=x_i|F_j) \right) P(F_j)$$

$$\begin{aligned}
 &= \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} x_i \cdot P(X = \bar{X} \text{ and } F_j \text{ occurs}) \\
 &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} x_i \cdot P(\bar{X} = x_i \text{ and } F_j \text{ occurs}) \\
 &\quad \swarrow \text{law of total probability} \\
 &= \sum_{i=1}^{\infty} x_i \cdot P(\bar{X} = x_i) \\
 &= \mathbb{E}(\bar{X}). \blacksquare
 \end{aligned}$$

So far, we have considered discrete r.v. \bar{X} . We can generalize $\mathbb{E}(\cdot)$ and $\text{Var}(\cdot)$ to continuous r.v. by replacing $P(\bar{X} = \dots)$ with $f(\dots)$ for the density function f :

\mathbb{R} -valued

Dfn: Let \bar{X} be a continuous r.v. with density $f(x)$. The expected value of \bar{X} is $\mathbb{E}(\bar{X}) = \int_{\mathbb{R}} x f(x) dx$.

Remark: This is clearly in analogy to the discrete case:

$$\sum_{i=1}^{\infty} x_i \cdot P(\bar{X} = x_i) \longleftrightarrow \int_{-\infty}^{\infty} x f(x) dx$$

Definition: Let \underline{X} be a \mathbb{R} -valued r.v. with density $f(x)$. The variance of \underline{X} is $\text{Var}(\underline{X}) = \int_{\mathbb{R}} [x - \mathbb{E}(X)]^2 \cdot f(x) dx$.

Ex: Let $f(x) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-[x-\mu]^2/(2\sigma^2)\right)$ be a Gaussian density. If $\underline{X} \sim N(\mu, \sigma^2)$ (i.e. \underline{X} has density $f(x)$), then:

$$\cdot \mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x) dx = \frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{\infty} x \exp\left([x-\mu]^2/(2\sigma^2)\right) dx$$

$$u = \frac{x-\mu}{\sqrt{2\sigma^2}} \quad = \frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{\infty} \left[u + \frac{2\sigma^2 \cdot u}{\sqrt{2\sigma^2}}\right] \exp\left(\cancel{-u^2}\right) \frac{1}{\sqrt{2\sigma^2 \cdot 2u}} du$$

$$du = dx \cdot \frac{1}{\sqrt{2\sigma^2}}$$

$$dx = \sqrt{2\sigma^2} \cdot du$$

$$x = \mu + \sqrt{2\sigma^2}u$$

$$= \cancel{\sqrt{2\pi}\sigma^2} \cdot \frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{\infty} \left[u + 2\sigma^2 u\right] \exp(-u^2) du$$

$$= \frac{1}{\sqrt{\pi}} \cdot \int_{-\infty}^{\infty} u \cdot \exp(-u^2) du + \frac{2\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u \exp(-u^2) du$$

$$= \frac{\mu}{\sqrt{\pi}} \cdot \int_{-\infty}^{\infty} \exp(-u^2) du \quad 0 \text{ by symmetry}$$

$$= \frac{\mu}{\sqrt{\pi}} \cdot \sqrt{\pi}$$

$\int_{-\infty}^{\infty} \exp(-u^2) du = \sqrt{\pi}$; check
 out Laplace's trick to use
 polar coordinates