Lecture 5 - Conditional Distribution of RV's & Inequalities

Conditional Probability for RV's

Recall $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$

A: event, B: event

TP Rule: $P(A) = \sum_i P(A \mid B_i) P(B_i)$

where B_i's form a partition of Ω

- Consider the discrete RV X with pmf $f_X(x) = P(X=x)$ and sample space Ω_X ($\equiv$ support)

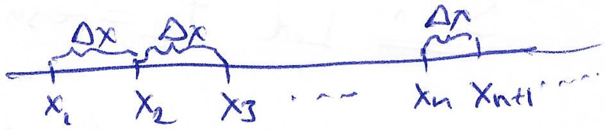
Then for any event A,

$$P(A) = \sum_{x \in \Omega_X} P(A \mid X=x) \underbrace{P(X=x)}_{f_X(x)}$$

"Total Prob. rule for discrete RV X"

- Consider the continuous RV. X with pdf $f_X(x)$ with support Ω_X ($\equiv$ sample space)

Partition Ω_X with x_i's : Ω_X



Then for any event A,

$$P(A) = \sum_i P(A \mid X \in (x_i, x_{i+1})) P(X \in (x_i, x_{i+1}))$$

$$= \sum_i P(A \mid X \in (x_i, x_i + \Delta x)) \underbrace{P(X \in (x_i, x_i+\Delta x))}_{\approx f_X(x_i)\Delta x}$$

As the partition gets finer and finer (i.e. as $\Delta x \to 0$)

$$P(A) = \int_{\Omega_X} P(A \mid \underbrace{X = x}_{\approx X \in (x, x+dx)}) f_X(x)\, dx$$

(TP Rule for cts RVs)

Similarly, for any r.v. Y,

$$E(Y) = \sum_{x \in \Omega_X} E(Y|X=x)\,P(X=x)$$

$$= \sum_{x \in \Omega_X} E(Y|X=x)\,f_X(x) \quad \text{if } X \text{ is discrete and } \Omega_X \text{ support of } X$$

"Law of Total Expectation for <u>Discrete</u> RV's"

Also

$$E(Y) = \int_{\Omega_X} E(Y|X=x)\,f_X(x)\,dx \quad \text{if } X \text{ is continuous and } \Omega_X \text{ is support of } X.$$

"Law of Total Expect. for <u>Continuous</u> RV's"

<u>Extension</u> $E[g(Y)] = \int_{\Omega_X} E[g(Y)|X=x]\,f_X(x)\,dx$

<u>Ex:</u> $X \sim \text{Geo}(p) \implies f_X(k) = P(X=k) = (1-p)^{k-1}p$

$$\boxed{\sum_{i=0}^{\infty} x^i = \frac{1}{1-x}}$$

Show that

$$E[X] = 1/p$$

"# of indep. flips to get first success"

<u>Sol'n:</u> Let I_1 = result of flip 1 $= \begin{cases} 1 & \text{wp } p \text{ if flip 1 is H (success)} \\ 0 & \text{wp } 1-p \text{ if flip 1 is T (failure)} \end{cases}$

Using LTE for X,

$$E[X] = E[X|I_1=1]\,P(I_1=1) + E[X|I_1=0]\,P(I_1=0)$$

Note that $E[X|I_1=1] = E[\#\text{ of flips to get } 1^{st} \text{ success} \mid \text{flip 1 is success}]$

$= E[1] = 1.$

$E[X|I_1=0] = E[\#\text{ of flips to get } 1^{st} \text{ success} \mid \text{flip 1 is failure}]$

$\Rightarrow E[X \mid I_1 = 0] = 1 + E[X]$ (will verify this!)

Thus, $E[X] = 1 \cdot p + (1 + E[X])(1-p)$

$= \cancel{p} + 1 - \cancel{p} + E[X](1-p)$

$\Rightarrow \cancel{E[X]} = 1 + \cancel{E[X]} - pE[X]$

$\Rightarrow pE[X] = 1 \Rightarrow E[X] = 1/p.$

$W =$ "$X \mid I_1 = 0$" $= \#$ of flips to get 1^{st} success | flip 1 is failure

$\underbrace{F}_{1} \mid \underbrace{\_\ \_\ \_\ \_\ \_\ \_\ \_}_{}$ start counting till first success $= V$

$\Rightarrow W = 1 + V$ where $V = \begin{cases} 1 & \text{wp } p \\ 2 & \text{wp } (1-p)p \\ 3 & \text{wp } (1-p)^2 p \\ \vdots & \vdots \\ k & \text{wp } (1-p)^{k-1} p \\ \vdots & \end{cases}$

$\Rightarrow P(V = k) = f_V(k) = (1-p)^{k-1} p, \quad k = 1, 2, \ldots$

$\Rightarrow V \overset{d}{=} X \sim \text{Geo}(p)$

so $E[V] = E[X]$

$\Rightarrow E[W] = E[X \mid I_1 = 0] = E[1 + V] = 1 + E[V]$

$= 1 + E[X].$ ■

Ex: Let X = time for a PhD student to graduate given the std has an MS degree

and Y = time for a PhD std to graduate given the std has a BS (but no MS) degree

Suppose that a professor believes

$X \sim \text{Uniform}(0,4)$ & $Y \sim \text{Uniform}(0,6)$ and $X \perp Y$

$P(Y<X) = ?$

Sol'n: (I)

$$P(Y<X) \stackrel{\text{TP rule}}{=} \int P(Y<X \mid X=x) f_X(x)\,dx$$

Detour: $U \sim \text{Uniform}(a,b) \Rightarrow f_U(x) = \begin{cases} \frac{1}{b-a}, & a<x<b \\ 0 & \text{o.w.} \end{cases}$

and $F_U(x) = \int_a^x \frac{1}{b-a}\,dx = \begin{cases} \frac{x-a}{b-a}, & a<x<b \\ 0 & x<a \\ 1 & x>b \end{cases}$

e.g. $\text{Uniform}(0,4) \Rightarrow a=0,\ b=4$

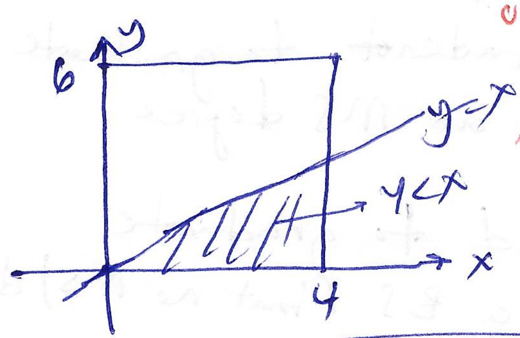
$$\Rightarrow P(Y<X) = \int_0^4 P(Y<x)\frac{1}{4}\,dx, \quad P(Y<x) = F_Y(x) = \frac{x}{6}$$

$$= \frac{1}{4}\int_0^4 \frac{x}{6}\,dx = \frac{1}{24}\int_0^4 x\,dx = \frac{1}{24}\left(\frac{x^2}{2}\Big|_0^4\right)$$

$$= \frac{1}{48}\cdot 16 = \frac{1}{3}$$

or (II) $P(Y<X) = \int_0^4\int_0^x f_{X,Y}(x,y)\,dy\,dx$

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \begin{cases} \frac{1}{4}\cdot\frac{1}{6} = \frac{1}{24}, & 0<x<4,\ 0<y<6 \\ 0 & \text{o.w.} \end{cases}$$



$$= \int_0^4\int_0^x \frac{1}{24}\,dy\,dx = \frac{1}{24}\int_0^4 \left(y\Big|_0^x\right)dx$$

$$= \frac{1}{24}\int_0^4 x\,dx = \frac{1}{24}\left(\frac{x^2}{2}\right)\Big|_0^4 = \frac{1}{3}$$

Exer: $P(X<Y) = ?$

Ex: $X_1, X_2, \ldots, X_n$ are iid RV's with mean μ and variance σ^2.

Let $S = \sum_{i=1}^{N} X_i$ where N is a RV $\perp X_i$ for all i.

Assume $N \geq 1$. Find $E[S]$.

Sol'n: $E[S] \overset{\text{TP rule}}{=} \sum_n E[S \mid N=n]\, P(N=n)$

$$= \sum_n E\left[\sum_{i=1}^{N} X_i \,\Big|\, N=n\right] P(N=n)$$

$$= \sum_n E\left[\sum_{i=1}^{n} X_i\right] P(N=n)$$

where $E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$ (LOE)

$$= \sum_{i=1}^{n} \mu \quad \text{(identical distr.)}$$

$$= n\mu$$

$\Rightarrow E[S] = \sum_n n\mu\, P(N=n) = \mu \underbrace{\sum_n n\, P(N=n)}_{E[N]}$

$$= \underline{\mu\, E[N]}$$

Probability Inequalities

Preliminaries

- $E[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx = \underbrace{\int_{-\infty}^{0} x f_X(x)\,dx}_{\leq 0} + \underbrace{\int_{0}^{\infty} x f_X(x)\,dx}_{\geq 0}$

$$\leq \int_0^\infty x f_X(x)\,dx$$

- If $X \geq 0$, $E[X] = \int_0^\infty x f_X(x)\,dx = \underbrace{\int_0^{10} x f_X(x)\,dx}_{\geq 0} + \int_{10}^{\infty} x f_X(x)\,dx \geq$

$$\int_{10}^{\infty} x f_X(x)\,dx \geq \underbrace{\int_{10}^{\infty} 10 f_X(x)\,dx}_{\text{since } x > 10} = 10 \int_{10}^{\infty} f_X(x)\,dx = 10\, P(X \geq 10)$$

$$\Rightarrow P(X \geq 10) \leq \frac{E[X]}{10}$$

[More generally, for any $t > 0$, $\boxed{P(X \geq t) \leq \frac{E[X]}{t}}$
(This is called Markov's Inequality).]

(Pf of MI) (for cts r.v.'s)

$$E[X] = \int_0^\infty x f_X(x)\,dx \geq \int_t^\infty x f_X(x)\,dx \geq \int_t^\infty t f_X(x)\,dx$$

$$= t \int_t^\infty f_X(x)\,dx = t\, P(X \geq t)$$

$$\Rightarrow P(X \geq t) \leq \frac{E[X]}{t}$$

- Chebyshev's Inequality:

Let X be a r.v. with mean μ and finite variance σ^2
$\Rightarrow E[X] = \mu$ and $Var(X) = \sigma^2 < \infty$

For $t > 0$, $P(|X - \mu| \geq t) \leq \frac{\sigma^2}{t^2} = \frac{E[(X-\mu)^2]}{t^2}$

<u>Pf:</u> Recall Markov's Inequality: For $X \geq 0$, $t > 0$

$$P(X \geq t) \leq \frac{E[X]}{t}.$$

So, we apply M.I. to $(X-\mu)^2$:

$$\rightarrow \quad P((X-\mu)^2 \geq t^2) \leq \frac{E[(X-\mu)^2]}{t^2} = \frac{\sigma^2}{t^2}$$

$$\Leftrightarrow \quad P(|X-\mu| \geq t) \leq \frac{\sigma^2}{t^2}$$

<u>Special Case:</u> Let $t = k\sigma$, $k > 0$

$$P(|X-\mu| \geq k\sigma) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$$

or equivalently

$$1 - P(|X-\mu| \leq k\sigma) \leq \frac{1}{k^2}$$

$$\Leftrightarrow \quad P(|X-\mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$$

["prob of X being within k std dev of mean μ is at least $1 - 1/k^2$"]

e.g. $P(|X-\mu| \leq \sqrt{2}\sigma) \geq 1 - \frac{1}{(\sqrt{2})^2} = 1 - \frac{1}{2} = \frac{1}{2} = 50\%$

$P(|X-\mu| \leq 2\sigma) \geq 1 - \frac{1}{4} = \frac{3}{4} = 75\%$

$P(|X-\mu| \leq 3\sigma) \geq 1 - \frac{1}{9} = \frac{8}{9} = 88.9\%$

- <u>Weak Law of Large Numbers (WLLN)</u>

$X_1, X_2, \ldots, X_n$ are iid RV's with $E[X_i] = \mu$ and variance $Var(X_i) = \sigma^2$ (possibly ∞)

Let $\bar{X}_n = \frac{\sum_{i=1}^n X_i}{n}$ (sample mean)

Then, for any $\varepsilon > 0$, $\lim_{n \to \infty} P(|\bar{X}_n - \mu| > \varepsilon) = 0$

Pf: $E[\bar{X}_n] = E\left[\frac{\sum X_i}{n}\right] = \frac{1}{n} E\left[\sum_{i=1}^{n} X_i\right]$

$= \frac{1}{n} \sum_{i=1}^{n} E[X_i]$ (LOE)

$= \frac{1}{n} \sum_{i=1}^{n} \mu$ (iid)

$= \frac{1}{n}(n\mu) = \mu$

$\mathrm{Var}(\bar{X}_n) = \mathrm{Var}\left(\frac{\sum X_i}{n}\right) = \frac{1}{n^2} \mathrm{Var}(\sum X_i)$

$= \frac{1}{n^2} \sum_{i=1}^{n} \mathrm{Var}(X_i)$ (LOV for indep. sums)

$= \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2$ (iid)

$= \frac{1}{n^2}(n\sigma^2) = \sigma^2/n$

Want to show

$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| > \varepsilon) = 0$$

By Chebyshev's Ineq,

$$P(|\bar{X}_n - \mu| > \varepsilon) \le \frac{\mathrm{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

$$\Rightarrow \lim_{n\to\infty} P(|\bar{X}_n - \mu| > \varepsilon) \le \lim_{n\to\infty} \frac{\sigma^2}{n\varepsilon^2} = 0$$

$$\Rightarrow \lim_{n\to\infty} P(|\bar{X}_n - \mu| > \varepsilon) = 0$$

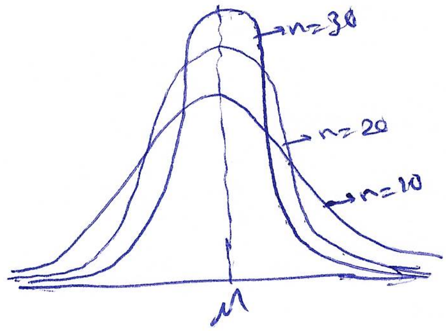
- **Central Limit Theorem (CLT)**

$X_1, X_2, \ldots, X_n$ are iid RV's with mean μ and variance $\sigma^2 < \infty$

Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ (sample mean).

Then $\bar{X}_n \overset{approx}{\sim} N(\mu, \sigma^2/n)$ for large n

or $$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0,1) \text{ as } n \to \infty$$



Recall $E[\bar{X}_n] = \mu$ & $Var(\bar{X}_n) = \sigma^2/n$.