

# Section 7: CLT, Joint Distributions, Law of Total Expectation (and bit of conditional distributions)

---

## Review of Main Concepts

- **Central Limit Theorem (CLT):** Let  $X_1, \dots, X_n$  be iid random variables with  $\mathbb{E}[X_i] = \mu$  and  $Var(X_i) = \sigma^2$ . Let  $X = \sum_{i=1}^n X_i$ , which has  $\mathbb{E}[X] = n\mu$  and  $Var(X) = n\sigma^2$ . Let  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ , which has  $\mathbb{E}[\bar{X}] = \mu$  and  $Var(\bar{X}) = \frac{\sigma^2}{n}$ .  $\bar{X}$  is called the *sample mean*. Then, as  $n \rightarrow \infty$ ,  $\bar{X}$  approaches the normal distribution  $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$ . Standardizing, this is equivalent to  $Y = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$  approaching  $\mathcal{N}(0, 1)$ . Similarly, as  $n \rightarrow \infty$ ,  $X$  approaches  $\mathcal{N}(n\mu, n\sigma^2)$  and  $Y' = \frac{X - n\mu}{\sigma\sqrt{n}}$  approaches  $\mathcal{N}(0, 1)$ .

It is no surprise that  $\bar{X}$  has mean  $\mu$  and variance  $\sigma^2/n$  – this can be done with simple calculations. The importance of the CLT is that, for large  $n$ , regardless of what distribution  $X_i$  comes from,  $\bar{X}$  is *approximately normally distributed with mean  $\mu$  and variance  $\sigma^2/n$* . Don't forget the continuity correction, only when  $X_1, \dots, X_n$  are discrete random variables.

- **Multivariate: Discrete to Continuous:**

	Discrete	Continuous
<b>Joint PMF/PDF</b>	$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$	$f_{X,Y}(x, y) \neq \mathbb{P}(X = x, Y = y)$
<b>Joint range/support</b> $\Omega_{X,Y}$	$\{(x, y) \in \Omega_X \times \Omega_Y : p_{X,Y}(x, y) > 0\}$	$\{(x, y) \in \Omega_X \times \Omega_Y : f_{X,Y}(x, y) > 0\}$
<b>Joint CDF</b>	$F_{X,Y}(x, y) = \sum_{t < x, s < y} p_{X,Y}(t, s)$	$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(t, s) ds dt$
<b>Normalization</b>	$\sum_{x,y} p_{X,Y}(x, y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$
<b>Marginal PMF/PDF</b>	$p_X(x) = \sum_y p_{X,Y}(x, y)$	$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$
<b>Expectation</b>	$\mathbb{E}[g(X, Y)] = \sum_{x,y} g(x, y) p_{X,Y}(x, y)$	$\mathbb{E}[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy$
<b>Independence</b> must have	$\forall x, y, p_{X,Y}(x, y) = p_X(x)p_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$	$\forall x, y, f_{X,Y}(x, y) = f_X(x)f_Y(y)$ $\Omega_{X,Y} = \Omega_X \times \Omega_Y$

- **Law of Total Probability (r.v. version):** If  $X$  is a discrete random variable, then

$$\mathbb{P}(A) = \sum_{x \in \Omega_X} \mathbb{P}(A|X = x)p_X(x) \quad \text{discrete } X$$

- **Law of Total Expectation (Event Version):** Let  $X$  be a discrete random variable, and let events  $A_1, \dots, A_n$  partition the sample space. Then,

$$\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X|A_i] \mathbb{P}(A_i)$$

- **Conditional Expectation:** See table. Note that linearity of expectation still applies to conditional expectation:  
 $\mathbb{E}[X + Y|A] = \mathbb{E}[X|A] + \mathbb{E}[Y|A]$

- **Law of Total Expectation (RV Version):** Suppose  $X$  and  $Y$  are random variables. Then,

$$\mathbb{E}[X] = \sum_y \mathbb{E}[X|Y = y] p_Y(y) \quad \text{discrete version.}$$

- **Conditional distributions**

	Discrete	Continuous
<b>Conditional PMF/PDF</b>	$p_{X Y}(x y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$	$f_{X Y}(x y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$
<b>Conditional Expectation</b>	$\mathbb{E}[X Y = y] = \sum_x x p_{X Y}(x y)$	$\mathbb{E}[X Y = y] = \int_{-\infty}^{\infty} x f_{X Y}(x y) dx$

- **Continuous Law of Total Probability:**

$$\mathbb{P}(A) = \int_{x \in \Omega_X} \mathbb{P}(A|X = x) f_X(x) dx$$

- **Continuous Law of Total Expectation:**

$$\mathbb{E}[X] = \int_{y \in \Omega_Y} \mathbb{E}[X|Y = y] f_Y(y) dy$$

## 1. Content Review

- (a) Select one: Suppose we have  $n$  independent and identically distributed random variables  $X_1, X_2, \dots, X_n$ , each with mean  $\mu$  and variance  $\sigma^2$ . Let  $X = \sum_{i=1}^n X_i$ . Then as  $n$  grows large, the Central Limit Theorem tells us that  $X$  behaves similarly to which normal distribution?

- $X \sim \mathcal{N}(n\mu, n\sigma^2)$
- $X \sim \mathcal{N}(\mu, n\sigma^2)$
- $X \sim \mathcal{N}(n\mu, \sigma^2)$
- $X \sim \mathcal{N}(n\mu, n^2\sigma^2)$

- (b) Suppose  $C$  and  $D$  are discrete random variables. Then  $\mathbb{E}[C|D = d] =$

- $\sum_d dp_{D|C}(d|c)$
- $\sum_c cp_{C|D}(c|d)$
- $\int_{-\infty}^{\infty} cf_{c|d} dx$
- $\frac{\mathbb{E}[C]}{\mathbb{E}[D]}$

- (c) Suppose  $X$  and  $Y$  are random variables and  $A$  is an event. Given that  $\mathbb{E}[X|A] = 4$  and  $\mathbb{E}[Y|A] = 10$ , what is  $\mathbb{E}[2X + Y/2|A]$ ?

- 14
- 18
- 9
- 13

### Central Limit Theorem Problems

The next few problems are CLT focused problems. Here's a general template for that! Sometimes we'll be trying to solve for the probability of something (e.g.,  $P(X \leq 10)$ ), and sometimes, we'll be trying to find a value of some parameter that will allow for the probability to be in a certain range (e.g.,  $P(X \leq 10) \leq 0.2$ ). Regardless, we still will want to apply CLT on  $X$ , and follow the same process (the only difference is that we may be solving for different things).

- (a) Setup the problem - write event you are interested in, in terms of sum of random variables. (what do we want to solve for/what is the probability we want to be true?)
  - Write the random variable we're interested in as a sum of i.i.d., random variables
  - Apply CLT to  $X = X_1 + X_2 + \dots + X_n$  (we can approximate  $X$  as a normal random variable  $Y \sim N(\mu, \sigma^2)$ )
  - Write the probability we're interested in
- (b) If the RVs are discrete, apply continuity correction.
- (c) Normalize RV to have mean 0 and standard deviation 1:  $Z = \frac{Y - \mu}{\sigma}$
- (d) Replace RV in probability expression with  $Z \sim N(0, 1)$
- (e) Write in terms of  $\Phi(z) = P(Z \leq z)$
- (f) Look up in the Phi table (or do a reverse Phi table lookup if we're looking for a value of  $z$  that gives us a certain probability)

## 2. Round off error

Let  $X$  be the sum of 100 real numbers, and let  $Y$  be the same sum, but with each number rounded to the nearest integer before summing. If the roundoff errors are independent and uniformly distributed between  $-0.5$  and  $0.5$ , what is the approximate probability that  $|X - Y| > 3$ ?

## 3. Bad Computer

Each day, the probability your computer crashes is 10%, independent of every other day. Suppose we want to evaluate the computer's performance over the next 100 days.

- (a) Let  $X$  be the number of crash-free days in the next 100 days. What distribution does  $X$  have? Identify  $\mathbb{E}[X]$  and  $Var(X)$  as well. Write an exact (possibly unsimplified) expression for  $\mathbb{P}(X \geq 87)$ .
- (b) Approximate the probability of at least 87 crash-free days out of the next 100 days using the Central Limit Theorem. Use continuity correction.

**Important:** continuity correction says that if we are using the normal distribution to approximate

$$\mathbb{P}\left(a \leq \sum_{i=1}^n X_i \leq b\right)$$

where  $a \leq b$  are integers and the  $X_i$ 's are i.i.d. **discrete** random variables, then, as our approximation, we should use

$$\mathbb{P}(a - 0.5 \leq Y \leq b + 0.5)$$

where  $Y$  is the appropriate normal distribution that  $\sum_{i=1}^n X_i$  converges to by the Central Limit Theorem.<sup>1</sup>

For more details see pages 209-210 in the book.

---

<sup>1</sup>

The intuition here is that, to avoid a mismatch between discrete distributions (whose range is a set of integers) and continuous distributions, we get a better approximation by imagining that a discrete random variable, say  $W$ , is a continuous distribution with density function

$$f_W(x) := p_W(i) \quad \text{when } i - 0.5 \leq x < i + 0.5 \text{ and } i \text{ integer}$$

## 4. Tweets

A prolific twitter user tweets approximately 350 tweets per week. Let's assume for simplicity that the tweets are independent, and each consists of a uniformly random number of characters between 10 and 140. (Note that this is a discrete uniform distribution.) Thus, the central limit theorem (CLT) implies that the number of characters tweeted by this user is approximately normal with an appropriate mean and variance. Assuming this normal approximation is correct, estimate the probability that this user tweets between 26,000 and 27,000 characters in a particular week. (This is a case where continuity correction will make virtually no difference in the answer, but you should still use it to get into the practice!).

## 5. Trinomial Distribution

A generalization of the Binomial model is when there is a sequence of  $n$  independent trials, but with three outcomes, where  $\mathbb{P}(\text{outcome } i) = p_i$  for  $i = 1, 2, 3$  and of course  $p_1 + p_2 + p_3 = 1$ . Let  $X_i$  be the number of times outcome  $i$  occurred for  $i = 1, 2, 3$ , where  $X_1 + X_2 + X_3 = n$ . Find the joint PMF  $p_{X_1, X_2, X_3}(x_1, x_2, x_3)$  and specify its value for all  $x_1, x_2, x_3 \in \mathbb{R}$ .

## 6. Do You “Urn” to Learn More About Probability?

Suppose that 3 balls are chosen without replacement from an urn consisting of 5 white and 8 red balls. Let  $X_i = 1$  if the  $i$ -th ball selected is white and let it be equal to 0 otherwise. Give the joint probability mass function of

- (a)  $X_1, X_2$
- (b)  $X_1, X_2, X_3$

## 7. Successes

Consider a sequence of independent Bernoulli trials, each of which is a success with probability  $p$ . Let  $X_1$  be the number of failures preceding the first success, and let  $X_2$  be the number of failures between the first 2 successes. Find the joint pmf of  $X_1$  and  $X_2$ . Write an expression for  $E[\sqrt{X_1 X_2}]$ . You can leave your answer in the form of a sum.

## 8. Continuous joint density

The joint density of  $X$  and  $Y$  is given by

$$f_{X,Y}(x,y) = \begin{cases} xe^{-(x+y)} & x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

and the joint density of  $W$  and  $V$  is given by

$$f_{W,V}(w,v) = \begin{cases} 2 & 0 < w < v, 0 < v < 1 \\ 0 & \text{otherwise.} \end{cases}$$

Are  $X$  and  $Y$  independent? Are  $W$  and  $V$  independent?

## 9. Trapped Miner

A miner is trapped in a mine containing 3 doors.

- $D_1$ : The 1<sup>st</sup> door leads to a tunnel that will take him to safety after 3 hours.
- $D_2$ : The 2<sup>nd</sup> door leads to a tunnel that returns him to the mine after 5 hours.
- $D_3$ : The 3<sup>rd</sup> door leads to a tunnel that returns him to the mine after a number of hours that is Binomial with parameters  $(12, \frac{1}{3})$ .

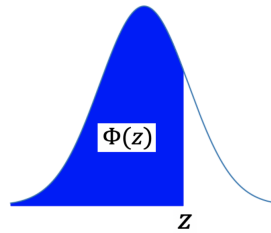
At all times, he is equally likely to choose any one of the doors. What is the expected number of hours for this miner to reach safety?

## 10. Lemonade Stand

Suppose I run a lemonade stand, which costs me \$100 a day to operate. I sell a drink of lemonade for \$20. Every person who walks by my stand either buys a drink or doesn't (no one buys more than one). If it is raining,  $n_1$  people walk by my stand, and each buys a drink independently with probability  $p_1$ . If it isn't raining,  $n_2$  people walk by my stand, and each buys a drink independently with probability  $p_2$ . It rains each day with probability  $p_3$ , independently of every other day. Let  $X$  be my profit over the next week. In terms of  $n_1, n_2, p_1, p_2$  and  $p_3$ , what is  $\mathbb{E}[X]$ ?

## 11. 3 points on a line

Three points  $X_1, X_2, X_3$  are selected at random on a line  $L$  (continuous independent uniform distributions). What is the probability that  $X_2$  lies between  $X_1$  and  $X_3$ ?



$\Phi$  Table:  $\mathbb{P}(Z \leq z)$  when  $Z \sim \mathcal{N}(0, 1)$

$z$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.6293	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.6591	0.66276	0.6664	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884	0.71226	0.71566	0.71904	0.7224
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.7549
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337	0.77637	0.77935	0.7823	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493	0.87698	0.879	0.881	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.9032	0.9049	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.9222	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.9452	0.9463	0.94738	0.94845	0.9495	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.9608	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.9732	0.97381	0.97441	0.975	0.97558	0.97615	0.9767
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.9803	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.983	0.98341	0.98382	0.98422	0.98461	0.985	0.98537	0.98574
2.2	0.9861	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.9884	0.9887	0.98899
2.3	0.98928	0.98956	0.98983	0.9901	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.9918	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.9943	0.99446	0.99461	0.99477	0.99492	0.99506	0.9952
2.6	0.99534	0.99547	0.9956	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.9972	0.99728	0.99736
2.8	0.99744	0.99752	0.9976	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.999