4.1 The Uniform Distribution

Def'n: A c.r.v. X has a continuous uniform distribution on $[a, b]$ when its *pdf* is

$$f(x) = \frac{1}{b-a} \qquad a \le x \le b$$

Also,

$$\mu = E(X) = \frac{b+a}{2} \quad \text{and} \quad \sigma^2 = V(X) = \frac{(b-a)^2}{12}$$

Ex4.1) Suppose X, the level of unbelievability at any point in a *Transformers* movie, is uniformly distributed between 100 and 1000.

a) What are the mean and variance?

b) What is the probability that the unbelievability level is above 500?

4.2 Exponential Distribution

Def'n: An exponential random variable represents the length of an interval from a certain point *until the next success* in a Poisson process. As such, the exponential distribution is like the geometric by looking for the prob. of the next success. A continuous random variable X has an exponential distribution with parameter λ (avg. # of successes per unit interval) when its *pdf* and *cdf* are

$$f(x) = \lambda e^{-\lambda x} \quad \text{and} \quad F(x) = 1 - e^{-\lambda x},\ x \ge 0,\ \lambda > 0$$

Also,

$$\mu = E(X) = \frac{1}{\lambda} \quad \text{and} \quad \sigma^2 = V(X) = \frac{1}{\lambda^2}$$

Like the Poisson distribution, use *consistent units* to define an exponential random variable X and parameter λ. Also, an exponential curve is skewed to the right, decreasing exponentially; as λ increases, the skewness of the curve increases. (See page 189.)

The exponential and Poisson distributions are *interchangeable*. Let an exponential random variable X with parameter λ denote the time until the next success. Let N denote the # of successes during a specific time interval x (i.e. N is a Poisson random variable with parameter λx). Then, the following relationship exists between X and N:

$$P(X > x) = P(N = 0) = \frac{e^{-\lambda x}(\lambda x)^0}{0!} = e^{-\lambda x}$$

Therefore, $F(x) = P(X \le x) = 1 - P(X > x) = 1 - e^{-\lambda x}$. By differentiating $F(x)$, *pdf* of X is

$$f(x) = \frac{dF(x)}{dx} = \lambda e^{-\lambda x},\ x \ge 0$$

→ Thus, we have an exponential distribution.

	Poisson	Exponential
Length of time interval	Constant	Variable (X)
# of successes	Variable (N)	Constant (1)

As seen in 3.2.1, the *lack of memory property* allows for exclusion of the history of previous outcomes such that the time can be reset. In other words,

$$P(X < t + \Delta t \mid X > t) = P(X < \Delta t)$$

(Proof)

$$P(X < t + \Delta t \mid X > t) = \frac{P(t < X < t + \Delta t)}{P(X > t)} = \frac{F(t + \Delta t) - F(t)}{1 - F(t)}$$

$$= \frac{1 - e^{-\lambda(t+\Delta t)} - (1 - e^{-\lambda t})}{1 - (1 - e^{-\lambda t})} = \frac{e^{-\lambda t} - e^{-\lambda(t+\Delta t)}}{e^{-\lambda t}} = 1 - e^{-\lambda \Delta t}$$

and $P(X < \Delta t) = F(\Delta t) = 1 - e^{-\lambda \Delta t}$

Thus, $P(X < t + \Delta t \mid X > t) = P(X < \Delta t)$.

Only the exponential distribution has this property among continuous distributions (as did *geometric* among the discrete distributions). This property does imply, though, that the value of λ does not change with X, it MUST be consistent.

Ex4.2) Presume a Poisson process linking every single game a team plays so that the time is, essentially, continuous. The # of goals by one team has a mean of 2.75 per "game" (or, 60 minutes).

a) Determine the *pdf* of the time (X; unit: 5 min) until the next goal.

b) Find the probability that there are no goals for at least 15 minutes by using both the exponential and Poisson distributions.

c) Find the mean and standard deviation of X.

5.1 Normal Distribution

Def'n: A c.r.v. X has a normal distribution with mean μ and variance σ^2 when its *pdf* is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad -\infty < x < \infty$$

The parameters μ (ranges from $-\infty$ to ∞) and σ^2 (> 0) are equal to $E(X)$ and $V(X)$, respectively. A normal distribution, denoted as $N(\mu, \sigma^2)$, is *symmetric* about μ and *bell-shaped*. The symmetry of a normal curve implies $P(X < \mu) = P(X > \mu) = 0.5$. The parameters affect the *center* and *shape* of the curve.

(diagrams drawn in class show each effect)

Probability of the Normal Distribution (a.k.a. the Empirical Rule):
The area beyond $\pm 3\sigma$ is quite small (less than 0.01). As in Figure 5.16, the two parameters combine for noteworthy properties:

1. 68% of observations lie within 1σ of μ.
2. 95% of observations lie within 2σ of μ.
3. 99.7% of observations lie within 3σ of μ.

This last property implies that 6σ is approximately the full width of a normal distribution.

Def'n: A normal r.v. Z has a standard normal distribution when $\mu = 0$ and $\sigma^2 = 1$. The *cdf* of Z is denoted by

$$\Phi(z) = P(Z \le z)$$

Tips & tricks:
- diagrams are helpful
- Complement: $P(Z \ge z) = P(Z > z) = 1 - P(Z \le z)$
- Symmetry: $P(Z \ge z) = P(Z \le -z)$
- $P(a \le Z \le b) = P(Z \le b) - P(Z \le a)$
- If $z > 0$, then $P(-z \le Z \le z) = 1 - 2P(Z \le -z)$

Ex5.1) *Examples with z-scores (finding prob.):*
a) $P(Z < -3.14) =$
b) $P(Z > 1.44) = 1 - P(Z < 1.44) =$
OR $P(Z > 1.44) = P(Z < -1.44) =$
c) $P(-3.14 \leq Z \leq 1.44) = P(Z \leq 1.44) - P(Z \leq -3.14) =$
d) $P(-2.00 \leq Z \leq 2.00) = 1 - 2P(Z \leq -2.00) =$

<u>5.1.3 Standardizing a Normal Distribution:</u>

$X \sim N(\mu, \sigma^2)$ and $Z \sim N(0, 1)$. What is Z? $Z = \dfrac{X - \mu}{\sigma}$, $z = \dfrac{x - \mu}{\sigma}$

$$P(X \leq x) \rightarrow P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = P(Z \leq z)$$

Ex5.2) *Examples with z-scores (standardizing):*
Find the following probabilities for $X \sim N(75, 6.5^2)$:

a) What is the probability of getting a value greater than 94.5?

b) What is the probability of getting a value between 71.75 and 84?

Identifying values:
Using the area under the curve, you can find appropriate z values; so, what are the corresponding x values?
$x = \mu + z\sigma$

Ex5.3) *Examples with z-scores (finding values):*
Use the same X as in Ex5.2) to answer the following:

a) What value denotes the top 5%?

b) What values bound the middle 70% of the data?

5.2 → revisits 2.6 for normal random variables

5.3.1 Normal Approximation to Binomial

When p is close to 0 or 1, the binomial distribution is quite skewed. In these cases, normal approximation is inappropriate. If, however, $np > 5$ AND $n(1 - p) > 5$, then the binomial distribution, $B(n, p)$, is approximately normal with $\mu = np$ and $\sigma^2 = np(1 - p)$. Thus, if $X \sim B(n, p)$, then

$$P(X \leq x) \approx \Phi\left(\frac{x + 0.5 - np}{\sqrt{np(1-p)}}\right) \qquad P(X \geq x) \approx 1 - \Phi\left(\frac{x - 0.5 - np}{\sqrt{np(1-p)}}\right)$$

Notes:

1. The 0.5 represents a "continuity correction" to improve the approximation.
2. The textbook likes the condition value of 5, but 15 is better and coincides better with later material, so exams will preferably avoid values between 5 and 15.

Ex5.4) Suppose that $X \sim B(100, 0.2)$. Find $P(X \leq 25)$ from a) the binomial distribution and b) if the approximation is appropriate, the approximate normal distribution.

Note also the following calculations:

Binomial	**Normal approximation**
$P(X \leq 15) = 0.1285$	$P(X \leq 15) = 0.1292$ (using 15 + 0.5)
$P(X \geq 25) = 1 - P(X \leq 24) = 0.1314$	$P(X \geq 25) = 0.1292$ (using 25 – 0.5)
$P(X \geq 15) = 1 - P(X \leq 14) = 0.9196$	$P(X \geq 15) = 0.9162$ (using 15 – 0.5)