Dr. Z.'s Probability Lecture 12 Handout: The Uniform and Normal Distribution

By Doron Zeilberger

Version of Oct. 27, 2017 (Thanks to Jinhua Xu)

Important Continuous Random Variable: The **Uniform Distribution** on the interval (a, b) has probability density function

$$f(x) = \begin{cases} \frac{1}{b-a} & if \quad a < x < b\,; \\ 0 & otherwise\,. \end{cases}$$

Important Facts: If X is the uniform distribution on the interval (a, b) then the expectation and variance are given by the formulas

$$E[X] = \frac{a+b}{2} \quad , \quad Var(X) = \frac{(b-a)^2}{12} \quad .$$

Problem 12.1: If X is uniformly distributed over $(10, 20)$, calculate the probability that (a) $X < 14$ (b) $X > 18$ (c) $11 < X < 16$.

Sol. to 12.1: The density function is

$$f(x) = \begin{cases} \frac{1}{10} & if \quad 10 < x < 20\,; \\ 0 & otherwise\,. \end{cases}$$

(a)

$$P(X < 14) = P(10 < X < 14) = \int_{10}^{14} f(x)\, dx = \int_{10}^{14} \frac{1}{10}\, dx = \frac{14-10}{10} = \frac{2}{5} \quad .$$

(b)

$$P(X > 18) = P(18 < X < 20) = \int_{18}^{20} f(x)\, dx = \int_{18}^{20} \frac{1}{10}\, dx = \frac{20-18}{10} = \frac{2}{10} = \frac{1}{5} \quad .$$

(c)

$$P(11 < X < 16) = P(11 < X < 16) = \int_{11}^{16} f(x)\, dx = \int_{11}^{116} \frac{1}{10}\, dx = \frac{16-11}{10} = \frac{1}{2} \quad .$$

Ans. to 12.1: (a) $\frac{2}{5}$ (b) $\frac{1}{5}$ (c) $\frac{1}{2}$.

Problem 12.2: Let X be uniformly distributed over (a, b) Derive (*from scratch*) (i) $E[X]$ (ii) $Var(X)$.

Sol. to 12.2:

$$E[X] = \mu = \int_a^b x \cdot \frac{1}{b-a}\,dx = \frac{1}{b-a}\int_a^b x\,dx = \frac{1}{b-a}\left(\frac{x^2}{2}\Big|_a^b\right)$$
$$= \frac{1}{b-a}\left(\frac{b^2}{2} - \frac{a^2}{2}\right) = \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2} \quad .$$

So $\mu = E[X] = \frac{a+b}{2}$. Next we derive the variance

$$Var(X) = E[(X-\mu)^2] = \int_a^b \frac{1}{b-a}(x-(a+b)/2)^2\,dx = \frac{1}{b-a}\int_a^b (x-(a+b)/2)^2\,dx \quad .$$

Let's make the **change of variable** $y = x - (a+b)/2$. Then $dx = dy$ and we get

$$Var(X) \;=\; \frac{1}{b-a}\int_{-(b-a)/2}^{(b-a)/2} y^2\,dy = \frac{1}{b-a}\left(\frac{y^3}{3}\Big|_{-(b-a)/2}^{(b-a)/2}\right)$$
$$= \frac{1}{b-a}\,\frac{((b-a)/2)^3 - (-(b-a)/2))^3}{3} = \frac{1}{b-a}\,\frac{2((b-a)/2)^3}{3} = \frac{(b-a)^2}{12} \quad .$$

Ans. to 12.2: $E[X] = \frac{a+b}{2}$ and $Var(X) = \frac{(b-a)^2}{12}$.

Problem 12.3: Trains arrive at a specified station at 20-minute intervals, starting at 8 AM. If a passenger arrives at a time that is uniformly distributed between 8 AM and 10 AM, what is the probability that he would have to wait

(a) Less than 13 minutes?

(b) between 5 and 11 minutes?

(c) between 5 and 11 minutes, if it is known that he had to wait less than 13 minutes.

Sol. to 12.3: He is equally likely to arrive at any twenty-minute interval between consecutive arrivals of the train. Let X be the arrival time after the previous train.

(a) If he has to wait less than 13 minutes it means that $X \geq 7$.

$$P(X \geq 7) = P(7 \leq X \leq 20) = \int_7^{20} \frac{1}{20}\,dx = \frac{20-7}{20} = \frac{13}{20} \quad .$$

(b) If he has to wait between 5 and 11 minutes it means that $X \leq 15$ and $X \geq 9$.

$$P(9 \leq X \leq 15) = \int_9^{15} \frac{1}{20}\,dx = \frac{15-9}{20} = \frac{6}{20} = \frac{3}{10} \quad .$$

(c)

$$P(9 \leq X \leq 15\,|\,7 \leq X \leq 20) = \frac{P(9 \leq X \leq 15)}{P(7 \leq X \leq 20)} = \frac{\frac{3}{10}}{\frac{13}{20}} = \frac{6}{13} \quad .$$

Ans. to 12.3: (a) $\frac{13}{20}$ (b) $\frac{3}{10}$ (c) $\frac{6}{13}$.

Problem 12.4: An insurance company issues policies covering damages to automobiles. The amount of damage is modeled by a uniform distribution on $[0, b]$.

The policy payout is subject to a deductible of $b/4$.

A policyholder experiences automobile damage.

Calculate the ratio of the standard deviation of the policy payout to the standard deviation of the amount of damage.

Sol. to 12.4: The variance of the amount of damage is easy, it is $\frac{b^2}{12}$.

The payout function is

$$P(x) = \begin{cases} 0 & if \quad 0 < x < \frac{b}{4}\,; \\ x - \frac{b}{4} & if \quad \frac{b}{4} < x < b\,. \end{cases}$$

The expected payout is

$$\int_{b/4}^{b} \frac{1}{b}(x - b/4)\,dx = \int_0^{3b/4} \frac{1}{b} x\,dx = \frac{1}{b}\,\frac{x^2}{2}\Big|_0^{3b/4} = \frac{1}{b}\,\frac{(3b/4)^2}{2} = \frac{9b}{32} \quad .$$

The expected square of the payout is

$$\int_{b/4}^{b} \frac{1}{b}(x - b/4)^2\,dx = \frac{1}{b}\int_0^{3b/4} x^2\,dx = \frac{1}{b}\,\frac{x^3}{3}\Big|_0^{3b/4} = \frac{1}{b}\,\frac{(3b/4)^3}{3} = \frac{9b^2}{64} \quad .$$

The variance of the payout is

$$Var(PayOut) = E[PayOut^2] - E[PayOut]^2 = \frac{9b^2}{64} - (\frac{9b}{32})^2 = \frac{63}{1024}\,b^2 \quad .$$

Hence the ratio of the variance of the policy payout to the variance of the damage is

$$\frac{\frac{63}{1024}\,b^2}{\frac{1}{12}b^2} = \frac{189}{256} \quad .$$

Finally, the ratio of the standard deviation of the policy payout to the standard deviation of the damage is the square-root of the above

$$\frac{3\sqrt{21}}{16} \quad .$$

Ans. to 12.4: The ratio of the standard deviation of the policy payout to the standard deviation of the amount of damage is $\frac{3\sqrt{21}}{16} = 0.8592329428\ldots$.

Most Important Continuous Random Variable: The **Normal Distribution** (aka **Gaussian Distribution**), with **parameters** μ and σ, denoted by $N(\mu, \sigma)$ is the continuous random variable whose density function is

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

Important Fact: The *expectation* of $N(\mu, \sigma)$ is μ and its standard deviation is σ, and hence its variance is σ^2.

Comment: Usually the "parameters" are μ, σ^2, but I prefer to use σ instead of σ^2. Of course, these are only conventions.

Very Important Special Case: The **Standard Normal Distribution**, Z, is $N(0, 1)$. In other words it is the continuous random variable whose pdf is

$$f(x) = \frac{1}{\sqrt{2\pi}}\ e^{-x^2/2} \quad , \quad -\infty < x < \infty \quad .$$

Very important fact: If X is a normal random variable with mean μ and standard-deviation σ (and hence variance σ^2), then

$$Z = \frac{X - \mu}{\sigma}$$

is the standard normal random variable.

The **cumulative distribution function** of Z is

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt \quad .$$

Unfortunately, there is no "nice" formula for it, so one has to use a computer (Maple has it) or a table. For example here:

`http://sites.math.rutgers.edu/~zeilberg/math477/Ztable.pdf` .

Problem 12.5: If X is a normal random variable with mean 2 and variance 9. find (a) $P(1 < X < 5)$; (b) $P(X > 1)$ (c) $P([X - 2| > 3)$.

Sol. to 12.5(a): Subtract $\mu = 2$ from both sides of $1 < X < 5$:

$$P(1 < X < 5) = P(1 - 2 < X - 2 < 5 - 2) = P(-1 < X - 2 < 3)$$

Next divide the inequality $-1 < X - 2 < 3$ by $\sigma = \sqrt{9} = 3$, getting that the desired probability is

$$P(-\frac{1}{3} < \frac{X-2}{3} < \frac{3}{3}) = P(-\frac{1}{3} < \frac{X-2}{3} < 1)$$

But $\frac{X-2}{3} = Z$, so we get, using the table

$$P(-\frac{1}{3} < Z < 1) = \Phi(1) - \Phi(-0.33) = 0.8413 - 0.3703 = 0.4710 \quad .$$

Sol. to 12.5(b): Subtract $\mu = 2$ from both sides of $1 < X < \infty$, and then divide by $\sigma = 3$

$$P(1 < X < \infty) = P(1-2 < X-2 < \infty) = P(-1/3 < \frac{X-2}{3} < \infty) = P(-1/3 < Z < \infty)$$

$$= \Phi(\infty) - \Phi(-0.33) = 1 - 0.3707 = 0.6293 \quad .$$

Sol. to 12.5(c):

$$P([X-2| > 3) = P(\frac{[X-2|}{3} > \frac{3}{3}) = P(|Z| > 1) = 1 - P(|Z| < 1) = 1 - P(-1 < Z < 1)$$

$$= 1 - 2P(0 < Z < 1) = 1 - 2(\Phi(1) - \Phi(0))$$

$$= 1 - 2\Phi(1) + 2\Phi(0) = 1 - 2\Phi(1) + 1 = 2(1 - \Phi(1)) = 2(1 - 0.8413) = 0.3174 \quad .$$

Important Fact: For large n, and p fixed, and not too small, the Binomial Distribution with parameters n and p, whose mass function is

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i} \quad ,$$

is approximated by the Normal Distribution with $\mu = np$ and $\sigma^2 = np(1-p)$.

To approximate $P(X = i)$ you pretend that it is $P(i - \frac{1}{2} < X < i + \frac{1}{2})$. To approximate $P(a \leq X \leq b)$ you pretend that it is $P(a - \frac{1}{2} < X < b + \frac{1}{2})$

Problem 12.6 Approximate the probability that if you toss a loaded coin, with $Pr(Head) = 1/3$, 1800 times

(a) you would get (strictly) more than 615 Heads;

(b) You would get (strictly) less than 550 Heads;

Sol. to 12.6: $\mu = 1800 \cdot \frac{1}{3} = 600$ and $\sigma^2 = \frac{1}{3}(1 - \frac{1}{3}) \cdot 1800 = 400$ so $\sigma = 20$.

(a)

$$P(615 < X) = P(616 \leq X) = P(615.5 \leq X) = P(X - 600 > 615.5 - 600) = P(X - 600 > 15.5) = P(\frac{X-600}{20} > \frac{15.}{20}$$

$$= P(\frac{X-600}{20} > 0.775) \quad .$$

This is approximated by $P(Z > 0.775) = 1 - \Phi(0.775) = 1 - 0.781 = 0.219$.

Note: The true value, according to Maple is 0.2188453507. Not bad!

Ans. to 12.6(a): The probability that you would get more than 615 Heads is *approximately* 0.219.

(b)

$$P(X < 550) = P(X < 549.5) = P(X-600 < 549.5-600) = P(X-600 < -50.5) = P(\frac{X-600}{20} < \frac{-50.5}{20})$$

$$= P(\frac{X-600}{20} < -2.525) \quad .$$

This is approximated by $P(Z < -2.525) = \Phi(-2.525) = 0.0058$.

Note: The true value is 0.005528537952. Not bad, but not great.

Problem 12.7: The working lifetime, in years, of a particular machine is normally distributed with mean 6 and variance 9. Calculate the 20-th percentile of the working lifetime, in years.

Sol. of 12.7: We want to find a number a such that

$$P(X < a) = 0.2 \quad .$$

Converting it to Z, using $\mu = 6$ and $\sigma = 3$,

$$P(X - 6 < a - 6) = 0.2 \quad .$$

$$P(\frac{X-6}{3} < \frac{a-6}{3}) = 0.2 \quad .$$

So we need

$$P(Z < \frac{a-6}{3}) = 0.2 \quad .$$

In other words

$$\Phi(\frac{a-6}{3}) = 0.2 \quad .$$

From the table

$$\Phi(-0.84) = 0.2 \quad .$$

Hence

$$\frac{a-6}{3} = -0.84 \quad ,$$

and so $a - 6 = 3 \cdot (-0.84) = -2.52$ and hence $a = 6 - 2.52 = 3.48$

Ans. to 12.7: The 20-th percentile of the working lifetime of the machine is 3.48 years.