STATISTICS FOR ENGINEERS
REVIEW PROBLEMS II

Here is a second helping of practice problems for the first examination.

Distributions

The probability mass functions (or probability density functions) are provided on their support. Outside the support, the probability is zero. I will not provide the support for the distributions. You should also be able to prove the expected value and variance formulae for all of the probability mass functions. For all of these, note that $q = 1 - p$.

Distribution	pmf (or pdf)	Expected Value	Variance
$K \sim \mathcal{U}(a, b)$	$f(k; a, b) = \frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2+1}{12}$
$K \sim Bern(p)$	$f(k; p) = p$	p	pq
$K \sim Bin(n, p)$	$f(k; n, p) = \binom{n}{k} p^k q^{n-k}$	np	npq
$K \sim Geo(p)$	$f(k; p) = pq^{k-1}$	$\frac{1}{p}$	$\frac{q}{p^2}$
$K \sim \mathcal{P}(\lambda)$	$f(k; \lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$	λ	λ
$X \sim \mathcal{U}(a, b)$	$f(x; a, b) = \frac{1}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$X \sim \mathcal{N}(\mu, \sigma^2)$	$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right]$	μ	σ^2
$X \sim Exp(\lambda)$	$f(x; \lambda) = \lambda e^{-\lambda x}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

Formulae

These formulae will be on the examination as they are on this page. It is up to you to know what each means and when to use each. Not all formulae are provided. I omitted those that are simple uses of elementary facts (linearity of expected value, e.g.).

$$\overline{X}; \bar{x}; \mu; \frac{\sum x_i}{n}; \sum x_i f(x_i); \int x f(x)\ \mathrm{d}x$$

$$s^2; \sigma^2; \frac{\sum (x_i - \bar{x})^2}{n-1}; \sum x_i^2 f(x_i) - \mathbb{E}[X]; \int x^2 f(x)\ \mathrm{d}x - \mathbb{E}[X]$$

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$$

$${}_nP_r = \frac{n!}{(n-r)!}$$

$${}_nC_r = \frac{n!}{r!(n-r)!}$$

$$\sigma^2_{c_1X_1 + c_2X_2 + \cdots + c_nX_n} = c_1^2\sigma^2_{X_1} + c_2^2\sigma^2_{X_2} + \cdots c_n^2\sigma^2_{X_n}$$

$$\sigma_U \sim \left|\frac{dU}{dX}\right| \sigma_X$$

$$\overline{X} \dot{\sim} \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

$$S_n \dot{\sim} \mathcal{N}(n\mu, n\sigma^2)$$

1. Sample Problems

Problem Some data

As a part of my job, my boss asked that I take a sample of 11 positioning tube and measure their lengths. I provide to you the lengths in the table below. He then wants me to find the arithmetic mean, the median, the variance, the standard deviation, the first and third quartiles, and the interquartile distance. I'm lazy. You do it for me.

13	15	11	11
12	15	14	11
12	13	11	

Solution: Arithmetic mean: $\bar{x} = 12.545(13)$; Median: $Q_2 = 12$; Variance: $s^2 = 2.4727(2.5)$; Standard deviation: $s = 1.5725(1.6)$; First quartile: $Q_1 = 11$; Third quartile: $Q_3 = 14$; Interquartile distance: $IQR = 3$. ◇

Problem A new distribution

The weight of a block, in kg, is distributed with pdf $f_X(x) = ce^{-2x}$, where x is the weight of the block, and x has support in $[0, 100]$. Find c. Find the expected value of the weight. Find the variance of the weight. Find the cumulative distribution function (cdf) of the weight. Find the median of the weight. Find the 95^{th}-percentile of the weight. Find the minimum weight. Find the maximum weight.

Solution: As f is a pdf, we know $\int_{-\infty}^{\infty} f_X(x)\ \mathrm{d}x = 1$. So, let us use that to determine c:

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} ce^{-2x}\ \mathrm{d}x \\ &= c\int_{0}^{100} e^{-2x}\ \mathrm{d}x \\ &= c\left[\frac{e^{-2x}}{-2}\right|_{0}^{100} \\ &= c\left(\frac{e^{-200}}{-2} - \frac{1}{-2}\right) \\ &\approx c\ 0.500 \end{aligned}$$

Thus, $c = 2$ and the pdf is $f_X(x) = 2e^{-2x}$.

To calculate the expected value, we use the formula: $\mu = \int_{-\infty}^{\infty} x f_X(x)\ \mathrm{d}x$:

$$\begin{aligned} \mu_1 &= \int_{-\infty}^{\infty} x2e^{-2x}\ \mathrm{d}x \\ &= \int_{0}^{100} x2e^{-2x}\ \mathrm{d}x \end{aligned}$$

Using integration by parts, we get that the expected value is $\mu_1 = 0.500$. To find the variance, we can either use the formula $\sigma^2 = \int_{-\infty}^{\infty} (x-\mu_1)^2 f_X(x)\ \mathrm{d}x$, or we can use $\sigma^2 = \mu_2 - \mu_1^2$, where $\mu_2 := \int_{-\infty}^{\infty} x^2 f_X(x)\ \mathrm{d}x$. Oddly enough, the second way is usually easier. You will need to do integration by parts, but you will find $\sigma^2 = 0.250$.

As for the cdf, it is yet another integration problem:

$$\begin{aligned} F_X(x) &:= \int_{-\infty}^{x} f_X(x)\ \mathrm{d}x \\ &= \int_{0}^{x} 2e^{-2x}\ \mathrm{d}x \\ \ldots &= 1 - 2e^{-2x} \end{aligned}$$

Now that we have the cdf, we can easily find the median (M)and the 95^{th} percentile (P_{95}):

$$\begin{aligned} 0.500 &= 1 - 2e^{-2M} \\ 0.950 &= 1 - 2e^{-2P_{95}} \end{aligned}$$

Thus, we get $M \approx 0.3466$ and $P_{95} \approx 1.498$.

The minimum weight is zero. The maximum weight is 100 pounds. ◇

Problem In the limit

A specific process produces paper with a thickness distributed $T \sim Exp(\lambda = 50)$ in inches. It is being used in a book of length 200 pages. What is the probability that a randomly selected book will be more than 5 inches thick?

Solution: This is a straight-forward application of the Central Limit Theorem. If T is the thickness of the book, we have $T \sim \mathcal{N}(\mu = 4; \sigma^2 = 2 \times 10^{-6})$. Calculating the z-score of the tested thickness (5 inches), we get $z = \frac{1}{0.001414} \approx 707.11$. Now, going to our Standard Normal table, we find that $\Phi(707.11) = 1$. Thus, we are all but guaranteed that the book will be thinner than 5 inches. $\diamond$

Problem A weighty issue

The average American weighs 178 pounds ($\sigma = 20$ pounds). Ten Americans get on an elevator that has a weight capacity of 1800 pounds. What is the probability that the weight limit is reached?

Solution: This is also a straight-forward application of the Central Limit Theorem. If W is the weight of those 10 people, then we know $W \sim \mathcal{N}(\mu = 1780, \sigma^2 = 40)$. Now, we calculate that the z-score is $z = \frac{20}{6.325} \approx 3.162$. Now, going to our Standard Normal tables, we determine that $\Phi(3.162) \approx 0.0008$. In other words, there is little to worry about, unless those ten Americans are not representative of Americans. $\diamond$