Chapter 7 – Continuous Random Variables

This chapter is the 'continuous analogue, of Chapter 4 (discrete). Whereas in Ch. 4, we spoke of pmf's (probability mass functions), here we deal with pdf's (probability *density functions*) since for continuous RV's, Pr(X = x) = 0 (*the Continuum Paradox*, p.196). Sums in Ch. 4 are replaced by integrals here.

The cdf of the continuous RV X is $F_X(x) = F(x) = \Pr(X \le x)$ as before. Now, however, the pdf is $f_X(x) = f(x) = F'(x)$. From the Fundamental Theorem of Calculus, we have

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

Example – the Continuous Uniform Distribution – Let X be uniformly distributed over the interval I = (a, b) where a < b. (Note in passing that we could have specified this interval as [a, b] – why?) For this distribution, the pdf is f(x) = 1/(b - a) for a<x<b, and the cdf is $F(x) = \int_a^x \frac{1}{b-a}dt = (x-a)/(b-a)$ for a<x<b. Graphs for a = 3 and b = 6 are given on pp. 196-7. In the general case, we write X ~ U(a, b).

→ The Properties of f and F on p.200 are important & should be noted.

Example – the Exponential Distribution – Whenever X has the pdf

$$f(x) = \frac{1}{\theta}e^{-x/\theta} \quad \text{for } x > 0$$

and unknown parameter θ > 0, we say that X has the Exponential(θ) distribution, and we write X ~ $\mathcal{E}(\theta)$. The corresponding cdf is then $F(x) = 1 - e^{-x/\theta}$, for x > 0.

Example 3. The Normal or Gaussian pdf with mean μ and variance σ^2 is given of p.203 top. We write $X \sim N(\mu,\sigma^2)$. The cdf has no closed form solution; cumulative probabilities are tabled (text front cover).

The *50th percentile* or *median* is the value of x* such that F(x*) = ½; similarly, for r between 0 and 100, the *rth percentile* is the value of x* such that F(x*) = r/100. The *mode* of a distribution is the value of x where the pdf f(x) reaches its maximum – obtained by differentiation.

For example, for the $\mathcal{E}(\theta)$ distribution, the 25th percentile is $\theta^*\ln(3/4)$, the median is $\theta^*\ln(2)$, and the 75th percentile is $\theta^*\ln(4)$.

Analogous to our Chapter 4 definitions, we define

- **the *mean* μ_X (or *expected value*) of X as $\mu_X = E(X) = \int x\, f(x)\, dx$**
- **the expected value of the function g(X) as $E(g(X)) = \int g(x)\, f(x)\, dx$**
- **the *variance* of X is $\sigma_X^2 = Var(X) = E\{(X - \mu)^2\} = \int (x - \mu)^2 f(x)\, dx$**
- **the shortcut formula still applies: $\sigma^2 = E(X^2) - \mu^2$**
- **if the RV X has mean μ and SD σ, $Z = (X - \mu)/\sigma$ is the *standardized random variable* corresponding to X**

(The standard deviation σ is the positive square root of the variance, and subscripts of X usually suppressed).

As in Chapter 4 (X is a RV, a and b are constants), we have:

1. **$E(aX + b) = a\,E(X) + b$**
2. **$Var(aX + b) = a^2\,Var(X)$**
3. **$SD(aX + b) = |a|\,SD(X)$**

<u>Theorem 7.4 – Chebyshev's Inequality</u> (p.213) states that for any RV X with finite mean μ and SD σ. Then for all r > 0,

$$Pr\{\mu - r\sigma < X < \mu + r\sigma\} > 1 - 1/r^2$$

This result is proven using Markov's Lemma, which is given on p.213. Chebyshev's rule is very strong since it applies to ALL random variables whereas the *empirical rule* applies to Normal RVs. See the comparisons in the center of p.213.

Some additional examples follow.

Example 4. The RV X has pdf $f(x) = 5x^4$ for $0 < x < 1$. The cdf is therefore $F(x) = x^5$ for $0 < x < 1$, so the median is $0.5^{1/5} = 0.8706$. The mean is $\mu = \int_0^1 5x^5\,dx = 5/6 = 0.8333$. Then, since $E(X^2) = \int_0^1 5x^6\,dx = 5/7$, $\sigma^2 = 5/7 - (5/6)^2 = 5/252$, and $\sigma = 0.1409$.

Example 5. The RV $X \sim \mathcal{E}(\theta)$, let's find the mean.

$$\mu = \int_0^\infty \frac{1}{\theta} x e^{-x/\theta}$$

$$= -xe^{-x/\theta}\Big|_0^\infty + \int_0^\infty \frac{1}{\theta} e^{-x/\theta} \quad \text{by integration by parts}$$

$$= -\theta e^{-x/\theta}\Big|_0^\infty = \theta$$

A similar calculation shows that $E(X^2) = 2\theta^2$, so $\sigma^2 = \theta^2$

Example 6. The RV Y has pdf $f(y) = 3(b-a)^{-3}(b-y)^2$ for $a < y < b$. Then $E(Y) = \frac{1}{4}(3a+b)$ – verify this. One application of this is that if X_1, X_2, and X_3 are independent uniform RV's on the interval (a,b), then $Y = \min\{X_1, X_2, X_3\}$ has this pdf and mean. So, when $a = 5$ and $b = 10$, then the expected value of this minimum is $25/4 = 6.25$.

Example 7. The RV Z has pdf $f(z) = 3(b-a)^{-3}(z-a)^2$ for $a < y < b$. Then $E(Y) = \frac{1}{4}(a+3b)$ – verify this. One application of this is that if X_1, X_2, and X_3 are independent uniform RV's on the interval (a,b), then $Z = \max\{X_1, X_2, X_3\}$ has this pdf and mean. So, when $a = 5$ and $b = 10$, then the expected value of this maximum is $35/4 = 8.75$.