Exercise 2

Two measurements *X* and *Y* are independently drawn from the same distribution with mean μ and variance σ^2, and a weighted sum $S = wX + (1-w)\cdot Y$ is computed, with $0 \le w \le 1$.

1) Find μ_S and σ_S^2.

2) Find the value of w that minimizes σ_S^2 and the minimum value for σ_S^2. Find an intuitive explanation for the findings.

3) Under the hypotheses of point 2), assuming that $\mu = 3, \sigma^2 = 8$ and that the distribution is symmetric around the mean value, compute $P\{S \le 7\}$.

4) Answer the above question again assuming that *X* and *Y* are normal.

5) Assume now that *X* and *Y* are *not* independent, and that $E[X\cdot Y] = \mu^2 + \Delta$. Answer again point 1). Is it possible that σ_S^2 decreases w.r.t. the previous case?

Solution

1) $\mu_S = E[S] = E[wX + (1-w)\cdot Y] = w\cdot E[X] + (1-w)\cdot E[Y] = w\cdot\mu + (1-w)\cdot\mu = \mu$

   Since *X* and *Y* are i.i.d.,

   $Var(S) = Var(wX + (1-w)\cdot Y) = Var(wX) + Var((1-w)\cdot Y) = \left[w^2 + (1-w)^2\right]\cdot\sigma^2$

   $= \left[2w^2 - 2w + 1\right]\cdot\sigma^2$

2) The required values are the coordinate of the vertex of parabola $y = 2w^2 - 2w + 1$, i.e. $w' = 1/2$, and $y' = 1/2$. Hence, the minimum value for σ_S^2 is $\sigma^2/2$. The intuitive explanation is that, when $w' = 1/2$, S is the average of *X* and *Y*. By the central limit theorem, the average of *n i.i.d.* random variables has a smaller variance than the individual variables.

3) We have $7 = \mu_S + 2\sigma_S$. Hence, by Tchebishev's inequality, it is $P\{|S - \mu_S| \ge k\cdot\sigma_S\} \le 1/k^2$, with $k = 2$. This means that $P\{S > 7\} + P\{S < -1\} = 1/4$. Therefore, since the RV is symmetric around the mean value, it is $P\{S > 7\} = 1/8$, hence $P\{S \le 7\} = 7/8$.

4) If *X* and *Y* are normal, it is $P\{S \le 7\} = P\{S \le \mu_S + 2\sigma_S\} = P\left\{\frac{S - \mu_S}{\sigma_S} \le 2\right\} = \Phi(2) = 0.9772$

5) μ_S stays the same, since the expectation is linear whether RVs are independent or not. The variance changes, and it incorporates the *covariance* between X and Y. More specifically, it is:

   $Var(S) = Var(wX + (1-w)\cdot Y)$

   $= Var(wX) + Var((1-w)\cdot Y) + 2\cdot Cov(wX, (1-w)\cdot Y)$

   $= w^2\cdot\sigma^2 + (1-w)^2\cdot\sigma^2 + 2\cdot w(1-w)\left[E[X\cdot Y] - \mu^2\right]$

   $= \left[2w^2 - 2w + 1\right]\cdot\sigma^2 + 2\cdot w(1-w)\cdot\Delta$

If $\Delta < 0$ (i.e. variables are negatively correlated), then the variance of S is actually smaller. This is because a large sample for X will be compensated by a smaller sample for Y and vice versa.